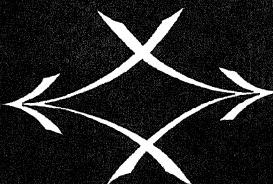


# PERSONALITY PATTERNS OF PSYCHIATRISTS



by ROBERT R. HOLT, Ph.D  
and LESTER LUBORSKY, Ph.D

*Introduction by*

ROBERT P. KNIGHT, M.D.

# PERSONALITY PATTERNS OF PSYCHIATRISTS

By  
ROBERT R. HOLT and  
LESTER LUBORSKY

What personal characteristics are most important in a good psychiatrist? How can a training institution scientifically evaluate a prospective psychiatrist's ability to master the lengthy, demanding preparatory work?

*Personality Patterns of Psychiatrists* is a pioneer effort to provide satisfactory answers to these searching questions. Based on Robert R. Holt's and Lester Luborsky's ten-year research study of 466 applicants for psychiatric training at The Menninger Foundation, this volume attempts for the first time to formulate systematic methods for testing, rating, and selection within an original theoretical framework.

In presenting their findings of the personal qualities most often associated with competence in psychiatrists, Drs. Holt and Luborsky employ illustrative material drawn from the case of one resident, whose record they analyze in detail.

Of *Personality Patterns of Psychiatrists*, Robert P. Knight, M.D., past president of the American Psychoanalytic Association, writes in his Foreword: "Every psychiatrist, psychoanalyst, and psychologist will find absorbing reading in this exceptionally clearly written study and will hail it as a greatly needed and genuine contribution to the literature. . . ."

40 00 00/00 00 00



3 1148 00274 6188

4/16 29 1979  
1

MA MAR 7 1981

MAN MAY 2 1979

MAN NOV 5 1981

132 H758p v.1  
Holt

61-12007

Personality patterns of  
psychiatrists

kansas city



public library

Kansas City, Missouri

Books will be issued only  
on presentation of library card.

Please report lost cards and  
change of residence promptly.  
Card holders are responsible for  
all books, records, films, pictures  
or other library materials  
checked out on their cards.

## *Personality Patterns of Psychiatrists*

**DATE DUE**

BY

*Robert R. Holt, Ph.D.*

AND

*Lester Luborsky, Ph.D.*

WITH THE COLLABORATION OF

*William R. Morrow, Ph.D.*

*David Rapaport, Ph.D.*

*Sibylle K. Escalona, Ph.D.*

*Menninger Clinic Monograph Series*

No. 13

PERSONALITY  
PATTERNS OF  
PSYCHIATRISTS

*A Study of Methods for  
Selecting Residents*

VOLUME I

PUBLISHERS    Basic Books, Inc.    NEW YORK

B  
B

© 1958 by Basic Books, Inc.  
Library of Congress Catalog Card No. 58-13159  
Manufactured in the United States of America  
Designed by Albert Margolies

*To Crusa and Ruth  
who gave up the time  
that this book took*



# FOREWORD

BY ROBERT P. KNIGHT, M.D.

WHAT PERSONAL QUALITIES SHOULD A YOUNG PHYSICIAN HAVE WHO wants to specialize in psychiatry? And how can the training institution to which he applies evaluate his potentialities for becoming a competent psychiatrist? The problem of selection of psychiatric residents is, of course, only one special division of the extremely important problem of vocational aptitude in the applicant and of personnel selection methods in the training or employing institution. In all such selection attempts it is essential to reduce to a minimum the waste of accepting and trying to train the misfits, and in psychiatry it might well be argued that selection is especially important. The cost of psychiatric training is high—estimated at between \$20,000 and \$35,000; the need for much larger numbers of psychiatrists is very great; the competence of each psychiatrist is of crucial importance to the many future patients he will have in his care; and the psychiatrists who do the training are supposed to know something about sizing up people. Yet, perhaps just because the shortage of physicians wishing to obtain psychiatric training has been so acute in the past, the selection procedures have been notably inferior in most institutions.

The Menninger School of Psychiatry, organized in 1945 through the inspiration and boundless energy of Dr. Karl Menninger, quickly became the largest training center for psychiatry in the world. Doctors returning from military service in World War II had been greatly impressed with both the need for and the therapeutic possibilities in psychiatry as a specialty, and they applied for training by the hundreds in the postwar years. It was my privilege to be a participant in the early years of this immense selection and training project,

and I well recall the many committee meetings, the discussions of rating scales and interview techniques, the never-ending series of interviews with applicants, and the Admissions Committee's deliberations over the interview reports and the psychological test reports of each applicant. Fortunately, the research-mindedness of the Menninger Foundation staff led to an early decision to take advantage of this golden opportunity by setting up a long-range research study in selection procedures. The authors of this volume were in charge of this research study, and they now report in thoroughgoing fashion on their lengthy and very largely pioneering efforts.

We have, in this volume, the definitive work on selection procedures for psychiatric residents. The authors have described the training setting and the teaching methods of the school and the comprehensive selection procedures used in assessing and choosing the trainees. They have placed this form of clinical selection in its proper perspective with regard to actuarial selection in personnel work, citing the relevant literature on the latter in detail. They have described with completeness their own research design, citing in considerable detail one trainee's case as an illustration. They have not been satisfied with reporting the actual results of the first five years, but have distilled out the principles and the practical advice to others that this rich study yielded. To enlighten the many psychiatric readers who may not have acquired a knowledge of statistical methods and terminology they have added a clear, concise section on the meaning of such methods and terms. A comprehensive bibliography and a complete index conclude this volume.

Every psychiatrist, psychoanalyst, and psychologist will find absorbing reading in this exceptionally clearly written study and will hail it as a greatly needed and genuine contribution to the literature on choosing and training young men for psychiatry.

## PREFACE

THIS BOOK TELLS THE STORY OF A TEN-YEAR EFFORT TO LEARN HOW residents were being selected for the Menninger School of Psychiatry and how the job could be done better. The last four of these years have been devoted to analyzing the results of the first six years' research and to writing this report.

The presentation in this volume is divided into four parts. In the first, we introduce the problem, our subjects, and the setting in which they were trained and the research was carried out. Part II gives the plans and main findings of the two major research designs we followed. The third part is devoted to the personality of the psychiatrist, with especial emphasis on qualities associated with competence. The final section of the book offers our conclusions and recommendations, for the selection of both psychiatric residents and psychoanalytic candidates.

Practical considerations have made it impossible for us to publish in a single volume all of the details of the research. The first volume contains a narrative account of the whole undertaking, written as nontechnically as the nature of the work permitted. It is self-contained, in the sense that the reader does not have to refer to Volume II in order to understand the procedure and the major findings. The second volume contains detailed accounts of the research methods and many tables of quantitative findings. It is arranged as a series of appendices to the chapters of Vol. I, and should be read in connection with these chapters at the points indicated in the text. References to tables, figures, or appendices numbered in arabic figures only refer to Vol. II; whenever the reference includes the roman numeral I (*e.g.*, Table I-4.1 in Chap. 4), the illustrative material is included in

this volume. Among the major items included in Vol. II are: detailed consideration of measures of competence in psychiatric work and full details of the methods we used in obtaining our criteria; samples of the manuals we developed for the scoring of interviews and projective tests for psychiatric selection; an extensive summary of expert opinion on the personality requisites for psychotherapy and psychoanalysis; and a full discussion of sources of error in our results — a section which may be of particular value to anyone who is undertaking research on the selection of professional people, particularly psychiatrists, psychoanalysts, or psychotherapists. It also contains a detailed, validated method of analyzing letters of recommendation and other credentials.

Many people have supported, encouraged, and aided us during these seven years. We are happy to name them here and acknowledge our debt of gratitude. The Veterans Administration financed the first five years, the New York Foundation the final two years of the research proper. The Menninger Foundation has supported the project since then and has throughout provided us with all manner of facilities and aids, from space and basic overhead to opportunities for many formal and informal consultations. Drs. Karl A. and William C. Menninger have not only given us generously of their time and advice but have smoothed the way for us to draw on all the resources of The Menninger Foundation and its School of Psychiatry. Without an institution of the size and outlook of the Menninger School of Psychiatry, no research of this kind would have been possible. Dr. Karl A. Menninger, Dr. Robert P. Knight, and Dr. David Rapaport conceived the project and planned its original design. Dr. Rapaport carried it single-handed throughout its first year and gave unstintingly of his time and energy to the project as long as he was the General Director of Research. From the time Dr. Rapaport left in 1948, through the succeeding four years of her directorship of the Department of Research, Dr. Sibylle K. Escalona took over from him the role of Responsible Investigator and gave us every intellectual and administrative support, as did Dr. Gardner Murphy, who became Director of Research in 1952. In this capacity and as chairman of the Menninger Clinic Monograph Committee, Dr. Murphy has given valuable advice on the analysis and presentation of our findings.

Early in 1948, after we had been on the job only half a year, Dr. William R. Morrow joined us as a third member of the team of co-directors for 18 crucial months in the development of the project.

He played an important role in the planning of the second experimental design and in helping carry it out. Mr. C. V. Ramana, a psychoanalyst from India in this country on a research fellowship, spent half time on the project from January to August 1948. Dr. Martin Mayman, of the Education Department in the Menninger Foundation, worked part time on some aspects of the project during 1946 and 1947. Mrs. Glenna Webster Tetzlaff worked as research assistant for the year starting July 1949, being succeeded for several months by Miss Irene Hollingsworth. Dr. Philip Holzman and Dr. Richard Siegal assisted with certain parts of the work a few months in 1949. Special mention should be made of the contribution made by Dr. Walter Kass, who spent a good part of 1950 and 1951 working on the development and cross-validation of a method for analyzing the Rorschach test for the selection of psychiatrists. Our statistical work for the first couple of years on the project was in the hands of Mr. Philip Chappell and Mr. James Knox. Miss Lolafaye Coyne has been a most stalwart assistant during the last eight years of the project. Without the high level of orderliness and accuracy that she introduced into the statistical work and tabulation of data we should hardly have been able to get the job done. We were fortunate in having as a research assistant for a year Mrs. Fae Kaufman Weiss, who had been with the Michigan project on the selection of clinical psychologists almost from the beginning. Finally, Mrs. Kathryn Brown was the mainstay (and often the only member) of our secretarial staff from November 1947 to September 1953, being assisted for a while first by Mrs. Twila Gough and then by Mrs. Korean Warner.

Many consultants have enriched our work with offerings from their wisdom and experience. From within the staff of the Menninger Foundation, we have been guided by Dr. Karl A. Menninger, Dr. William C. Menninger, Dr. Robert P. Knight, Dr. Jan Frank, Dr. Merton M. Gill, Dr. Robert L. Worthington, Dr. Lewis L. Robbins, Dr. Rudolf Ekstein, and Dr. Herbert Modlin. At a critical early stage in the research, Dr. James G. Miller (then Chief Psychologist in the Central Office of the Veterans Administration) and Dr. John Whitehorn, of the Johns Hopkins University and Phipps Clinic, helped to set our feet in a constructive direction. Dr. Donald Marquis of the University of Michigan gave useful counsel about the experimental design. Dr. Talcott Parsons advised and stimulated us in thinking about the sociological aspects of our project. Dr. Frederick Hacker, of the Hacker Clinic, advised us about our functional analysis of the

psychiatric resident's job. During the time that he was first Clinical Director and then Manager of Winter V.A. Hospital, Dr. Jesse F. Casey helped us with a host of problems and in a number of capacities. Our special gratitude goes to that large group of faculty members of the Menninger School of Psychiatry whom we called collectively "the supervisors": Service Chiefs and their assistants, consultants, attending physicians, and all the others who gave hundreds of hours to provide thoughtful and insightful criterion evaluations of the residents' work. The members of the various Admissions Committees since the beginning of the School have earned our thanks by their uncomplaining and faithful performance of special tasks for the sake of the research. Dr. Bert E. Boothe, formerly Director of Professional Education at Winter V.A. Hospital, supplied many kinds of information and assistance. The Registrar of the Menninger School of Psychiatry, Mr. George Hedstrom, has cheerfully complied with hundreds of requests.

The following persons have read major portions of the manuscript, and have given us detailed and helpful critique and suggestions: Drs. Lee J. Cronbach, Rudolf Ekstein, George S. Klein, Robert P. Knight, Karl A. Menninger, Gardner Murphy, Paul W. Pruyser, David Rapaport, Norman Reider, Lewis L. Robbins, William Rottersman, and Roy Schafer.

For all of this assistance, support and encouragement, we are very deeply grateful. Most of all, we want to express our thanks to the 466 physicians who have been our cooperative and very helpful subjects in an undertaking that was bound to be personally inconvenient and often threatening to them. It has been a privilege to work with them.

*New York  
Topeka  
September 1958*

ROBERT R. HOLT  
LESTER LUBORSKY

# CONTENTS

*Foreword* by Robert P. Knight, M.D. vii  
*Preface* ix

## P A R T I

### *The Setting and the Subjects*

1. Introduction to the Problem	3
2. Historical Background	7
3. The Menninger School of Psychiatry	18
4. The Entering Resident	33
5. Changes in the Residents During Psychiatric Education	52
6. After Residency	70

## P A R T II

### *The Research: Its Methods and Findings*

7. Theoretical Issues in Selection Research	79
8. The Problem of the Criterion	108
9. The Success of Selection by Rule of Thumb	128
10. Plans for a New Experimental Design	137
11. A Specimen Analysis of Predictions on One Subject	159
12. Principal Results of the Second Design	198
13. Validities of "Liking" and Some Additional Predictors	216
14. A Backward Look at Some Predictive Errors	234

P A R T   I I I

*The Personality of the Psychiatrist*

15. A Survey of Expert Opinion on Personality Requisites for Psychiatry	255
16. Personality Differences Between Better and Poorer Residents	272

P A R T   I V

*Conclusions and Implications*

17. Recommendations	297
18. Applications to the Selection of Candidates for Psychoanalytic Training	337
19. Overview	347

<i>Appendix: Statistical Note</i>	359
-----------------------------------	-----

<i>Bibliographic References</i>	365
---------------------------------	-----

<i>Index</i>	377
--------------	-----

P A R T I

THE SETTING AND  
THE SUBJECTS



## INTRODUCTION TO THE PROBLEM

WHO SHOULD BE CHOSEN FOR TRAINING IN PSYCHIATRY? AMERICAN psychiatrists with years of experience in selecting and training physicians for this specialty asked themselves this question with a new urgency at the end of World War II, for there was an unprecedented rush of doctors applying for residency training in psychiatry. Training facilities were being expanded, but could not keep up with the demand. Those who were responsible for schools of psychiatry were aware of the need for more knowledge about how to select the most promising prospects for psychiatric education—knowledge that could be provided only by research.

Besides the immediate practical need, there were other reasons that it was important to do the best possible job of choosing residents. The training institutions feel a duty to society and to the profession to supply the best psychiatrists in the largest possible numbers. They also have a heavy responsibility to the physicians applying, because undertaking psychiatric training is a commitment to a new identity and way of life, involving a major investment by the resident, the institution and the public in this expensive form of education. If a man is taken into training who will not complete it, there can result much loss, unhappiness, and sometimes even

#### 4 The Setting and the Subjects

danger to the resident and to the patients in his care. If a man who might become a good psychiatrist is rejected for training, a critically needed man is lost to a profession.

When the Menninger School of Psychiatry started in 1946, over six hundred physicians applied for the hundred residency positions that were to be available the first year. The leaders of the School worked out, and later published in an issue of the *Bulletin of the Menninger Clinic* (14:92-101, 1950), the standards by which they would judge the potential psychiatrist. It seemed a unique opportunity to start a research project into methods of selection, for among the essentials for such a project are large numbers of applicants and psychiatrists-in-training in a single institution.

Two of The Menninger Foundation's principal psychiatrists, Drs. Karl A. Menninger and Robert P. Knight, met with Dr. David Rapaport, director of the department of research, and devised the first research plans. The main goals were to study the accuracy of the School's selection methods and to determine whether any improvements could be made. In brief, the method of selection that was adopted was to have three staff psychiatrists interview each applicant and a psychologist administer a battery of diagnostic psychological tests. The opinions of these four were reported to an Admissions Committee who reviewed this information, together with reports of physical examination, medical school transcript, and letters of recommendation, and decided to accept or reject the applicant.

The Menninger School of Psychiatry was only one of many institutions for residency training, and all faced similar problems. An informal survey of opinion and practices concerning selection (made some time after the research had begun) showed that the same needs were felt in most of the large psychiatric training centers in the United States. Dr. Karl Menninger wrote to twenty-four chiefs of service asking them to describe and discuss their methods of selecting residents. Specifically, he asked:

1. Do you require personal interviews?
2. Do you pay any attention to letters of recommendation?
3. Do you use psychological tests?
4. Upon what qualities or evidences do you rely most heavily for your decisions?
5. How well satisfied are you with your decisions a year or two later (how many men have you had to drop or wish you could drop, in terms of percentages)?

Answers came back from seventeen training centers \* altogether, giving the best available picture of selection procedures and practices used in American psychiatry in 1950. *Personal interviews* were required by ten of the centers, were used whenever possible by four more, and were not required at all by the remaining three. These last relied on an evaluation of documentary materials submitted by the applicant, in one case according to the requirements of a state Civil Service regulation. The consensus of thirteen answers on *letters of recommendation* was that they were considered useful only when the writer was personally or professionally well known to someone at the training center; three institutions replied simply that letters of recommendation were taken seriously, and another that they were not. *Psychological tests* were used routinely at four centers; of these, one used only the Strong Vocational Interest Blank, another the Minnesota Multiphasic Inventory and the Miller Analogies Test, and two (including the Menninger School of Psychiatry) a battery of tests including the Rorschach and Thematic Apperception Test. At three centers, tests were used occasionally when indicated; at another, residents were tested *after* being accepted, but the tests were given scant attention unless the results were extremely bad. In the other nine institutions, tests were not used at all.

About their satisfaction with selection methods used, there was little consensus. Eleven centers reported quantitative estimates of the proportion of unsatisfactory results, ranging from 4 per cent to 33 per cent of all those they accepted. A majority of the eleven reported that they were dissatisfied with 20 per cent or more of the applicants accepted. From five institutions the replies mentioned no figures but evaluated their results as follows: "quite well satisfied," "relatively well satisfied," "moderately satisfied," "difficult to answer . . . we have had some quite troublesome problems with some of them but in several cases these men have proved to be unusually good psychiatrists," "very unhappy about our present technique." All told, about a third of the centers reported dissatisfaction with

\* The centers which contributed to this informal survey are the following: Boston Psychopathic Hospital, Brooklyn State Hospital, Elgin State Hospital, Institute of the Pennsylvania Hospital, Langley Porter Clinic, Menninger Foundation, Michael Reese Hospital-Associated Facilities, New York Hospital, New York State Psychiatric Institute, Norristown State Hospital, North Little Rock V.A. Hospital, University of Cincinnati, University of Colorado, University of Michigan, University of Minnesota, University of Texas, Washington University, St. Louis.

## *6 The Setting and the Subjects*

the methods used. Actually, however, no one *knew* how good a job of selection was being done anywhere.

The research reported in this book was undertaken to study the process of assessing applicants for psychiatric training and predicting their future competence, to learn how well it operated, what its strong and its weak points were, and, if possible, to point the way toward better methods.

## 2

### HISTORICAL BACKGROUND

THE PROBLEMS FACING THE ADMISSIONS COMMITTEE IN 1945 AND 1946 were not unique to Topeka or even to psychiatry. For the first time, selection researches were getting under way in clinical psychology (Kelly and Fiske, 1951), psychiatric social work (Berengarten, 1951), and psychoanalysis (H. Klein, 1953), all of which professions were also being deluged with applicants for training. What was happening nationally to bring about these sudden developments in the mental health professions?

Starting with the immediate cause, many opportunities were opening up as a consequence of a huge public demand for psychiatric and allied services. Awareness of mental illness, recognition of it in many of its forms and stages, realization that much could be done for it, and a clamor for such services, all emerged increasingly after World War II.\*

World War II was an important landmark in the development of American psychiatry. The fact that 38 per cent of rejections for

\* The developments to be described were, of course, the outgrowth of trends that had been developing for at least forty years. For further historical background, see standard works on the history of psychiatry, such as: J. K. Hall, *et al.* (1944), *One hundred years of American psychiatry*; Albert Deutsch (1937), *The mentally ill in America*; Wm. C. Menninger (1948b), *Psychiatry, its evolution and present status*. See also Appendix 2.1, Vol. II.

## *8 The Setting and the Subjects*

service were for psychiatric reasons (including 14 per cent for mental deficiency) shocked the American public. Despite this psychiatric screening, the tremendous pressures of war caused psychiatric crack-ups among large numbers of men: about 1 soldier in 13 had at least one neuropsychiatric admission to an Army hospital, and 40 per cent of all disability discharges were for such reasons. The total manpower loss from mental disorders, over two and a half million men, could have manned 177 Army infantry divisions. All of these figures were widely publicized and undoubtedly helped bring about an increasing awareness that mental illness is the nation's foremost health problem. With the gradual increase in sophistication about mental illness, it was somewhat easier in discussing war neuroses to speak frankly in terms of psychological breakdowns instead of in evasive terms, such as "shell-shock," that were used in the first great war, and psychiatrists were able to explain such breakdowns in ways that laymen could understand. Even more important, they were able to help, overworked and understaffed though they were, for many valuable treatment methods had been invented.

Moreover, medical men often found that an assignment to duty on psychiatric wards in the hospitals of the armed services was not the tedious and meaningless experience that their courses in psychiatry in medical school had led them to believe. Something was known about mental illness, they discovered, and much could be done about it. At the least, they had first-hand contact with the fascinating phenomena of mental illness. The high incidence of psychiatric casualties meant that many physicians with no previous experience or interest in psychiatric problems had to cope with them. In addition, great numbers of them had short but intensive training courses in psychiatry, many conducted by the best teachers in the profession. The result was a tremendous awakening of interest in psychiatry on the part of physicians, many of whom had settled into other types of medical careers before the war but who left the services determined to get further psychiatric training.

Similar developments were taking place in the other mental health professions. The war gave, if anything, a greater boost to the development of clinical psychology and began a movement to create a new kind of mental health worker, the psychiatric aide (Hall *et al.*, 1952). Although the emphasis in this book must necessarily be on psychiatry and psychiatrists, we want to mention here the important part played by these related professions, together with psychiatric

social work and psychiatric nursing, in caring for the mentally ill and advancing the growth of the mental health movement.

#### REFORM IN VA AND STATE HOSPITALS

The war created a whole new population of patients: war veterans whose psychological scars caused many immediate, and many more delayed, disturbances demanding psychiatric care. Psychiatric care for veterans was available through the Veterans Administration (VA)—at least, on paper. Unfortunately, the policies of the VA concerning medical care had caused its hospitals (or “facilities,” as they were officially called) to have a poor reputation among physicians. The great forward strides of the thirties in psychiatry had hardly roused an echo in these backward institutions. And even if they had been offering the best of care, they had available only a fraction of the needed beds.

A vigorous movement began in the public press, calling attention to the conditions in VA and state hospitals for emotionally disturbed persons. Overcrowding was the rule; in 1941, available hospital beds for emotionally disturbed people equaled only 63 per cent of the total estimated need. Through their professional organization, psychiatrists had adopted standards for their hospitals, including ratios of personnel to patients, that would have allowed at least humane custody if they had ever been enforced. But the association did nothing effective to put the standards into effect, and public support for the budgets they would have required was lacking. Not a single governmentally supported hospital in the country met the minimal standards set by the American Psychiatric Association. The understaffed institutions were operating on grossly insufficient budgets. When ordinary medical hospitals needed \$6 or more a day to care for each patient, some state institutions were budgeting *total* care of patients at less than 60¢ each—an amount not great enough to provide a minimal diet even if it were all spent for food (Deutsch, 1948). Antiquated rules kept the few psychiatrists in such institutions so swamped with paper work and the care of emergency physical illness that they had no time for direct psychiatric work with patients. Untrained and ignorant attendants, frightened by behavior they could not understand, in some cases terrorized patients by brutal treatment.

In 1945, a congressional committee began an investigation of

## 10 *The Setting and the Subjects*

the VA hospitals, producing more publicity and public education. Psychiatrists and other leaders in the mental hygiene movement spread the word that mental illness was not always incurable, that the practice of providing only borderline custodial care was in fact an enormously expensive waste of human resources and public funds, even in short-run terms of dollars and cents. A young man could be committed to a hospital as "insane" and languish there for decades with bed and board at state expense, although much less money might be needed to cure him. In addition, the VA system had its own special weaknesses. Its hospitals had often been badly located, in relatively inaccessible places. They were in the charge of non-medical managers who had no conception of using the hospital as a means to create a therapeutic milieu. And they had held themselves separate from affiliations with medical schools and other educational institutions (see Appendix 2.2).\*

This last fact was in some ways the most damaging of all. An active teaching program in a hospital attracts good staff men and keeps them abreast of the latest developments in therapeutic and diagnostic science. Teaching programs provide hospitals with a great deal of medical manpower, too, for the interns and residents who are trained in a hospital do much of its routine medical work while receiving in return their postgraduate training and experience.

The new administration in the VA, led by Generals Bradley and Hawley, made many reforms. Attracted by its new and progressive goals, many capable physicians who had previously been uninterested in working for the VA now joined the staff of its hospitals and set about providing the good medical care that had been promised to veterans. Since most of the beds in VA hospitals were, and still are, psychiatric beds (as is true of all hospitals in this country), the level of psychiatric care began to improve. Active building programs were instituted to relieve overcrowding and provide space for the many new patients who could be expected; from August 1945 through December 1947, 22 new neuropsychiatric hospitals with 7673 beds, were opened. If there were to be doctors and other professional staff to man these hospitals, the VA itself would have to train them. The old policy of aloofness was reversed, and residency programs in special fields of medicine and dentistry were

\* This and all succeeding references to appendices, tables, etc., numbered in arabic figures only refer to Vol. II. Illustrative material in this volume carries the identifying roman numeral I (e.g., Table I-4.1).

begun as quickly as possible, as well as training programs in clinical psychology and other fields.

The state hospitals were slower to catch up, but, as the conditions in them were publicized in popular magazines and newspapers, and by such books as Mary Jane Ward's *The Snake Pit* (1946), Harold Maine's *If a Man Be Mad* (1947), and Albert Deutsch's *The Shame of the States* (1948), they too began to improve. Larger appropriations meant better salaries for more and better personnel. Teaching programs were set up in state hospitals. Yet, as psychiatric institutional jobs improved in all ways and increased in numbers, so did opportunities in private practice arise. In 1948, only 60 per cent of U.S. psychiatrists were working in tax-supported hospitals—a decline of 20 per cent in 20 years. (For statistics on private practice see Appendix 2.1.)

These changes within the profession were both a cause and a result of a shift in the public's conception of mental illness. A central fact was the penetration of certain ideas into the awareness first of psychiatrists and then of the general population—ideas that mental illness did not consist of the major psychoses alone, conceived of as dread and hopeless scourges, but included a great variety of less severe conditions, and that these pathological changes in feelings, behavior, and thinking can be caused by life experience in any one of us and can be helped by the experience of psychotherapy as well as by such traditionally medical agents as drugs. But, of course, effective treatment required trained personnel.

The result of all these component trends was a huge increase in the demands for psychiatric care shortly after the war. There were fewer than three hundred psychoanalysts in the United States in 1946 and only about four thousand psychiatrists (see Fig. I-2.1), many of them old-line institutional men without training or experience in psychotherapy, and many inexperienced even with somatic treatments. Estimated minimal requirements called for at least three times as many psychiatrists as existed. A pitifully small group facing the Herculean labors they had now to do! (Further data on psychiatric shortages are given in Appendix 2.1.)

Obviously, the demand was going to continue to mount. To increase sharply the number of trained psychiatrists, facilities for psychiatric education had to expand, and they did so all over the country. Before we go on to the part of this expansion that took place in Topeka, the founding of the Menninger School of Psychi-

## 12 *The Setting and the Subjects*

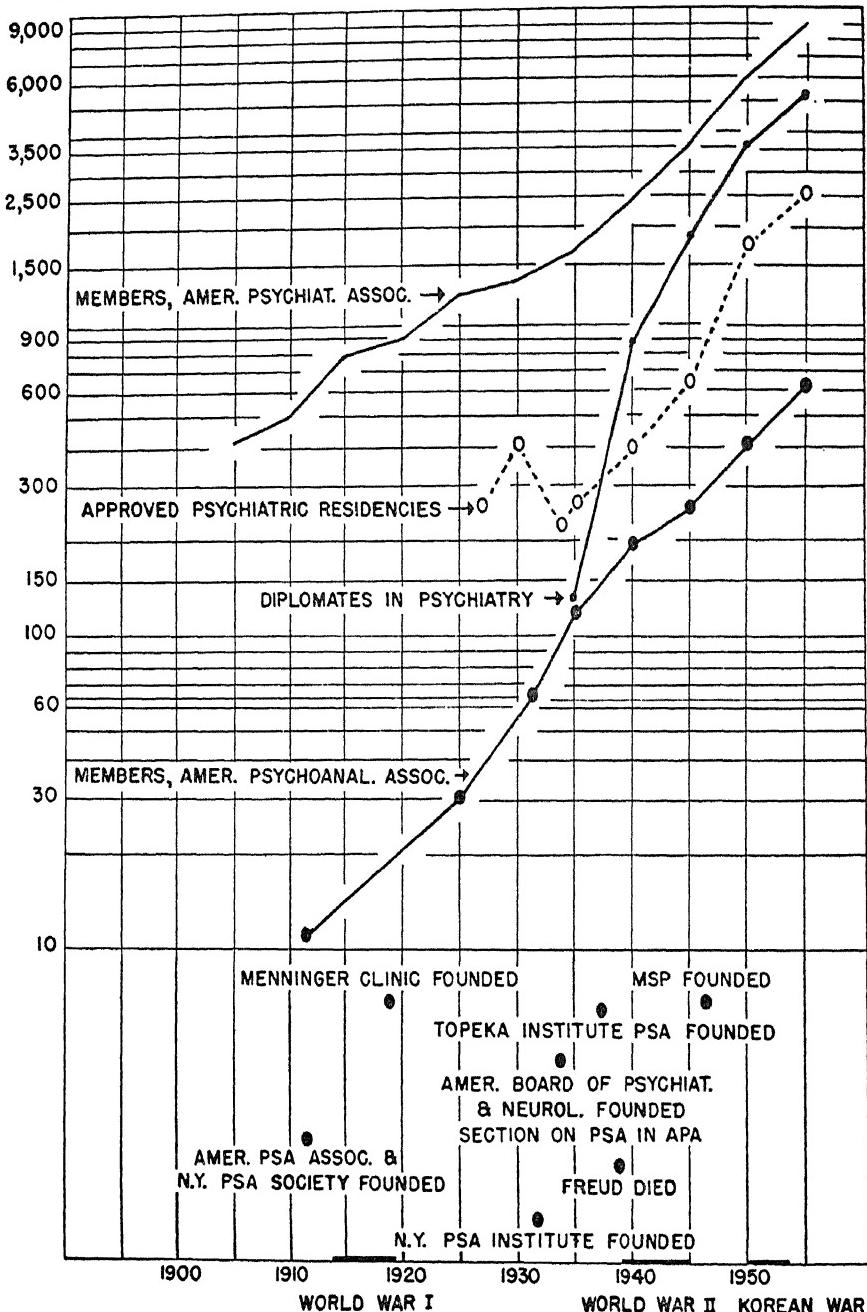


FIG. I-2.1. SOME TRENDS IN THE RECENT HISTORY OF AMERICAN PSYCHIATRY

try (Chap. 3), let us take a quick glance at the state of American psychiatric education at the war's end.

#### PSYCHIATRIC EDUCATION IN MEDICAL SCHOOLS

The first step in training a psychiatrist is taken in medical school, since he must first be a physician. In 1940, all American medical schools included some instruction in psychiatry—about 2 per cent of the total hours in the curriculum—and 87 per cent of them offered it in the preclinical years as well as in the later ones. Since the war, the amount of psychiatry in the curriculum has increased, as have the numbers of medical students.

The serious shortage of physicians after the war, plus the greatly increased number of applicants (many of whom were veterans), stimulated a tremendous program of expansion for the nation's medical schools. At a cost of over \$300,000,000, major efforts were made to increase facilities for medical education without lowering standards. For the five years before World War II, there was an annual enrollment of about 21,500 medical students; in 1951, the number had increased to more than 27,000—better than a 25 per cent increase. Twice as much money was being spent by medical schools (partly because of inflation) (Anderson *et al.*, 1952). Yet even with this huge effort, medical schools in the United States were turning out little more than 6000 M.D.'s a year, which does not seem a large figure when one considers that the population was growing by about 2,675,000 each year and that death, retirement, and other causes were removing approximately 1000 physicians from active service annually. (Further details in Appendix 2.1.)

Even if every one of these graduates had been psychiatrically trained, it would have taken years to provide the nation with the thousands of psychiatrists who were urgently needed. Of course, needs in other medical fields were increasing, too, and an indispensable condition for a man to become a psychiatrist is that he want to be one.

It is a painfully expensive and slow procedure to expand facilities for medical education. At the same time, we must remember that the men who are teaching psychiatry in medical schools today were themselves trained in medicine and psychiatry quite a few years ago. As a result, some of them have conservative (sometimes downright anachronistic) views about what psychiatry is and what it can do.

## 14 *The Setting and the Subjects*

Some concrete data backing up the last point were provided by representatives from the Menninger School of Psychiatry who visited a number of medical schools around the country and talked to graduating seniors about their plans (Hall, 1949). They found that many students held the following misconceptions: that psychiatry offers no private practice but only institutional jobs; that there are no effective forms of psychiatric therapy; that psychiatry as a discipline is vague, mystical, and unscientific; that the incomes of psychiatrists are very low; that psychiatry is cut off from other forms of medical practice; and many others. None of these ideas was completely true any more, yet many young physicians who might otherwise have wanted to seek training in psychiatry were discouraged because of such misconceptions, which they learned in medical school or before. Furthermore, most of the groups contacted were dissatisfied with their undergraduate medical courses in psychiatry (see also Chap. 4).

But there are welcome signs of a growing movement for more and better psychiatric teaching in medical schools. During the last few years, two national conferences have been held to consider this very problem. Psychoanalysts in increasing numbers are holding appointments in medical schools, and the dynamic as opposed to the descriptive viewpoint is constantly gaining ground. In a number of progressive medical schools, psychiatrists are taking a leading role in medical education. Being concerned with the patient as a whole, as a *person* who is in trouble rather than as merely the vehicle of an interesting liver or fractured femur, psychiatry is in a crucial position to help reorganize medical training along organismic lines. One sign of this growth is that in 1955 the faculties of United States medical schools had more unfilled positions in psychiatry than in but two other clinical departments; and in all twenty of the medical departments, only in internal medicine, surgery, and pathology were there more vacancies than in psychiatry (Turner *et al.*, 1955).

The outlook for the future, then, is somewhat better: graduates of medical schools will increasingly be better oriented to an understanding of the kind of profession that psychiatry is today, and they will be better grounded in some of its rudiments and better able to help the psychiatric aspects of the conditions their patients bring to them. Nevertheless, problems of serious mental illness will continue to be mainly the province of those medical men who have

specialized in psychiatry with the thoroughness that requires post-graduate education.

#### PSYCHIATRIC EDUCATION AFTER MEDICAL SCHOOL

To provide the educational facilities to meet the challenge and opportunity created by scientific advance, the medical profession in this country has established a series of specialty boards. Each such board, made up of a group of leaders in a medical specialty, offers a diploma certifying that its holder is a recognized specialist in that field. The board prescribes the standards of education and experience that the applicant must meet in order to obtain this diploma. The educational standards, in turn, describe what is known as residency training, and there are provisions for the periodic inspection and approval of residency training programs.

The American Board of Psychiatry and Neurology, Incorporated, was founded in 1934 to perform the services just mentioned for psychiatry and neurology. The Board has fixed the requirements that a physician must complete an approved internship of one year, three years of training in an approved residency, and two additional years of experience before he is eligible to be examined. After his eligibility is established, the candidate is given an oral examination. Only after he passes it does he become a Diplomate and a recognized specialist in psychiatry.

Before the war, psychiatric residencies were available primarily in university hospitals and a few state mental hospitals. The educational experience was in many ways more like that of an apprentice in a medieval guild than like that of a student in a modern university. The resident worked with a "Board man," observing him whenever possible, doing his routine work, going on ward rounds with him, and helping him in research if he happened to be doing any. Few institutions had more than a handful of residencies. If there were interns, the first-year or junior resident often had teaching responsibilities toward them, and he was equally likely to do some teaching of nurses or other groups in training in the same hospital. As he advanced to intermediate and senior resident status in his second and third years, he was given more responsibility both for the care of patients and for the training of less experienced residents. A certain amount of reading would be prescribed by the staff men

## 16 *The Setting and the Subjects*

with whom he worked, and there might be occasional lectures and seminars, but the emphasis was primarily clinical rather than didactic. Not many specialists could be trained in this way, and not many were.

The number of residencies in psychiatry offered in any one year fluctuated between approximately 200 and 400 from 1927 through 1940, the number of programs slowly growing to 111, while the profession itself increased its numbers by 50 per cent (see Figure I-2.1). In 1945, the number of programs being given at least temporary approval by the American Medical Association was growing so fast that it is difficult to give a figure representative of the whole year, but in October there were 632 residencies in 139 institutions. The flood of physicians seeking training after wartime service had not yet descended in full force, and only 34 per cent of these residencies were as yet filled.

An astonishing growth took place during the first five years of our research. By 1950, there were no less than 231 approved residency programs in psychiatry, offering 1754 appointments! Even though 80 per cent of them were filled, there were more vacant residencies in psychiatry than in any other specialty—344 of them. The growth during the subsequent five years has not been quite so rapid, but more residencies were filled in 1955 than the total number offered in 1950.

When considering these figures, particularly those on unfilled residencies, remember that any residency is no better than the teachers and the clinical institution involved. The best-rated programs were getting plenty of applicants, even though others had many vacancies. Another cause of unfilled residencies, especially since 1950, has been the requirements of the armed services, for the "doctor draft" hit especially at the age range when most physicians seek specialty training.

An important recent event in psychiatric history was the passage of the National Mental Health Act in 1946, which created the National Institute of Mental Health. This Institute has greatly stimulated and supported the growth of psychiatry and related fields through its subsidies for training programs in psychiatry, clinical psychology, psychiatric social work, and psychiatric nursing. During its first five years, the Institute awarded 430 stipends to resident psychiatrists to help make advanced training available to them (in

addition to direct grants to several new programs to help them pay their teachers).

Finally, after the war, money began to become available for psychiatric research. The Research Branch of the National Institute of Mental Health has been an important and steady source of funds for psychiatric and allied research. And the VA set up a contract research program, a grant from which supported for its first five years the research reported in this book.

## 3

THE MENNINGER SCHOOL  
OF PSYCHIATRY

THE DEVELOPMENT OF THE MENNINGER SCHOOL OF PSYCHIATRY WAS part of the national postwar expansion in the facilities for psychiatric residency training. Its history starts with the founding of the Menninger Clinic,\* a group practice set up by Dr. Karl A. Menninger with his father, Dr. Charles F. Menninger, just after World War I. In 1925 the Southard School was founded as a residential center for the treatment of disturbed children, and not long afterwards the hospital buildings of the Menninger Sanitarium were built. As early as 1929 the Clinic and Sanitarium had begun training a few psychiatrists.

In 1941 the Menninger Foundation was established as a non-profit institution, devoted to psychiatric treatment, education, and research. Eventually the assets of the Clinic, the Sanitarium, and the Southard School were transferred to the Foundation.

During this time and the wartime period that followed, the

\* The history of the Menninger Clinic, stressing the role of the Menningers themselves and particularly the father, Dr. C. F. Menninger, has been told in a lively (if not always accurate) way by Walker Winslow (1956). The same book contains an anecdotal account of the Menninger School of Psychiatry, based in large part upon the first draft of this chapter. It supplements the present account by emphasizing biographical and anecdotal materials, and gives a particularly good picture of the atmosphere of the School's exciting first year.

program of psychiatric training grew. When Winter General Hospital was taken over by the Veterans Administration, The Menninger Foundation agreed, in 1945, to manage the psychiatric training program in that hospital, with the unprecedented number of 100 residencies. Thus the Menninger School of Psychiatry came into being.

#### WINTER HOSPITAL FROM JANUARY 1946 TO JULY 1948

The total bed capacity of Winter Hospital at the time the Veterans Administration took it over was 2276. A number of buildings were remodeled by the Veterans Administration for other purposes, and the bed capacity was finally reset at 1400: 900 for psychiatric patients, 100 for neurological, 250 for general medical, and 150 for surgical patients.

Shortly after the first of January 1946, the first VA patients arrived, and 27 residents matriculated. The influx of patients and staff was rapid for the next few months, leveling off at nearly 1400 patients and about 1700 employees of all kinds. Twenty-one more residents came in April, and a group of 44 in July.

The atmosphere of the exciting (and sometimes confusing) period of organization had its effects upon the first classes of residents studied in this research and upon those who trained them, so we must describe it at least briefly. It was both exhilarating and stormy. There was sometimes conflict between the advanced philosophy underlying the Foundation's aspirations regarding training and administration and the more staid outlook of those practiced in the older forms of VA hospital operation and sensitive to the limitations of government machinery. Dr. Karl Menninger had agreed to serve as manager of the hospital during its formative stage; he remained in this position for two and a half years, introducing many innovations. Drs. Hawley and Blain time and again proved their willingness to support policies and procedures unprecedented in VA annals.

Recruitment of staff members was difficult at first because psychiatric training programs were expanding almost everywhere; but once the reputation of the hospital began to spread it became easier, and several experienced physicians were brought in during the first year and a half. First-rate staffs in the adjunctive therapies, social work, clinical psychology, and other nonmedical professions were gathered, in rather quick order. On his return from the direction of psychiatry

## 20 *The Setting and the Subjects*

in the Army, Dr. William Menninger took an active part in the School of Psychiatry, as well as becoming General Secretary of The Menninger Foundation.

Although it was always true that most of its patients were psychiatric, Winter VA Hospital was and has remained a *general* hospital. Its Medical Service, under Dr. Rudolph Chess, had a staff of eight internists and a residency program in internal medicine. The Neurological Service had a distinguished neurosurgeon and clinical neurologist. By the middle of 1948, several physicians with permanent status in the Department of Medicine and Surgery completed the psychiatric training they had come for and became increasingly useful in helping to train others. A number of the most experienced residents from the first groups, who had now completed all requirements of the American Board of Psychiatry and Neurology, accepted staff positions as service chiefs, rounding out a complete and competent staff.

The concept of the "psychiatric team" played a leading part at Winter Hospital. The manager tried to convey this idea directly to all employees. In a speech to a group of newly hired personnel—and new groups arrived weekly at first—he would say, for example: "We don't feel that doctors cure people. We feel that everybody in the hospital works to cure people—the doctors, the nurses, the social workers, the ward aides, the janitors, the people who tend the fires, the people who keep the books. *Everybody.*"

Everybody helped, and almost everybody was drawn into the expanding educational programs as teacher or student—often in both roles. The huge national shortage of trained personnel meant that the VA was going to have to undertake its own training programs on a major scale if it was to fulfill the country's promises to the veterans. One result was that many different kinds of training programs were concentrated together at Winter Hospital, because other trainees than those in psychiatry needed the opportunity to work in a first-class clinical setting, and because some of the teaching could be made available to clinical psychologists and psychiatric social workers. A number of the psychiatric residents helped with a program of in-service training for 330 psychiatric aides, whose training gave them not only clinical skills but self-respect as a responsible professional group.

The team concept and its extension to include everyone who worked in the hospital was a vital part of the *therapeutic milieu* that

was created at Winter. The adjunctive therapies were used intensively in Medical Rehabilitation, as many as 80 per cent of the patients participating daily in some part of the full program of activities. (By contrast, there are few mental hospitals in which more than half the patients regularly participate in such programs.) Working through the media of arts and crafts, music, classroom education in academic subjects, machine shops where patients could learn useful manual skills, picnics, parties, participation in sports and dances—through a busy program of purposeful and constructive activities—trained therapists developed relationships that were appropriate to the patient's needs, as prescribed by the psychiatrist.

The results soon began to attract attention, first over the nation and ultimately all over the world. The discharge rate of patients started at a high level; in the tenth month of the hospital's existence it had turned over 50 per cent of its running capacity of 1400 patients. (In interpreting these figures, one should remember that Winter was a *general* hospital.) In September 1946, the VA Advisory Committee on Psychiatry brought 40 consultants, including leaders of the professions included in the psychiatric team, to a three-day meeting at Winter Hospital. It was made plain that Winter was being developed as a pilot plant, a model for other VA hospitals and psychiatric residency programs. (At this time there were 16 training centers in the VA, with a total of 210 psychiatric residents; 102 of them were in Topeka.)

But of course there were also many inadequacies during this exciting period. The physical plant was poorly adapted to good psychiatric practice; there were vexing lacks of supplies and of clerical help; and it did not prove easy to modernize the VA ways of doing things in all aspects. Looking back on these days from the perspective of a few years later, one staff member wrote, "The work of the total staff was coordinated rather in spirit than by tested administrative procedures or by efficient habits of communication." The system of quarterly rotation of residents through the services was often hard on the patients, some of whom had as many as four doctors a year. During the first year, there were not enough experienced staff men to supervise and teach the Fellows. It was fortunate for the hospital that the residents who came during the first year were with few exceptions men who had had considerable psychiatric experience in the Army; psychiatric patients were no novelty to them, and they did not need the basic procedural orientation and emotional weathering

## 22 *The Setting and the Subjects*

that were to take up a good deal of the time of later classes. These first classes of residents played a major role in making the hospital run, day in and day out.

### WINTER HOSPITAL FROM 1948 TO 1951

If the early residents differed from the later ones in the experience they brought with them to the hospital, the hospital differed in the atmosphere it offered to the earliest and to the later classes of residents. The initial, developmental phase of the hospital's history came to an end about the middle of 1948. As the confusion and difficulties of the first years subsided, so did the initial exhilaration. But perhaps the two most important changes in the hospital after 1948 had to do with its being specifically a veterans' institution. It achieved a special place among the VA's teaching hospitals; perhaps it is fair to say, a pre-eminent place. Nevertheless, standard VA Rules and Procedures made a strong comeback after the time had passed when it was necessary for red tape to be cut in order to put the whole enterprise into motion. The new Rules and Procedures were by no means identical, however, with what they had been before the war; many changes resulted from innovations that began in Winter Hospital itself.

The other principal development stems from the inevitable dwindling of the early intense public interest in veterans' needs. As the war receded into history, Congress became less eager to give the VA all the funds it requested, and there was a series of crises over threatened (and to some extent realized) budget cuts. It was impossible to maintain the original ratio of personnel to patients, and inevitably the program suffered. These cutbacks dealt heavy blows to morale among staff and residents alike. Several physicians, social workers, and other professional staff members resigned. Many persons felt that the VA's future plans could not be certain; graduating residents began to refuse positions with the VA, and the number of accepted applicants for residency who changed their plans and went elsewhere increased also. Under the managership of Dr. J. F. Casey until June 1952, and after that of Dr. R. C. Anderson, the hospital nevertheless continued essentially its original methods of treating patients, remaining an active treatment center and never filling up with a residue of "untreatable" chronic patients.

THE MENNINGER SCHOOL OF PSYCHIATRY: CURRICULUM

When it began, The Menninger Foundation School of Psychiatry was in a number of ways unlike any other psychiatric training program in the world. Since that time, there has been some convergence between it and other large residency programs. An obvious difference still remains: it is about three times as large as the one with the next most numerous group of residents. Peak enrollment in the Menninger School of Psychiatry during its first five years was 111, at the end of July 1947; average enrollment was a little under one hundred.\*

Basically, the Menninger School of Psychiatry saw itself as different because it was organized as a school, whereas most other psychiatric training programs were primarily accessory to hospitals. The School was founded as a separate educational institution, with a curriculum, and with a faculty made up of staff members from The Menninger Foundation, the University of Kansas Medical School, Winter VA Hospital, and other organizations. The students were not called residents but Fellows (following the custom at many university training centers). And during the first phase, there was a good deal of emphasis on didactic and theoretical teaching.

The Menninger Clinic had long been regarded as a psychoanalytically oriented institution. Its senior staff members had made and were making contributions to the theoretical as well as to the clinical literature of psychoanalysis, and they believed firmly that this complex intellectual system needed to be taught to residents. Not every psychiatrist could learn the technique of psychoanalytic treatment, but it was thought important that every psychiatrist taught in the Menninger School of Psychiatry should learn about psychotherapy and enough of psychoanalytic theory to understand his patients' difficulties and to acquire some insight into the therapeutic process. At the same time, the curriculum and the clinical instruction were designed to teach the necessary procedures of psychiatric diagnosis and a wide range of therapeutic resources.

As time permitted, after the first great burst of organization, the School experimented almost constantly with its didactic curriculum

\* During the years since the termination of the selection research, it has again enlarged; in 1956, there were 135 residents.

## 24 *The Setting and the Subjects*

and clinical teaching methods. Moreover, with succeeding classes the educational needs of the residents changed. More and more residents came to the School directly from internships, with usually far less clinical experience of a psychiatric nature than the earliest residents, who had had military service. The use of lectures to large groups declined, and increasing reliance came to be placed on teaching small groups on the services, with the focus on day-to-day problems and specific patients rather than on general principles and theories. In October 1946, the Director of Professional Education wrote in the monthly report of the School that "the greater part of the curriculum is the clinical instruction." Ultimately, out of approximately 600 faculty hours expended per month, the School was scheduling only about 10 per cent for lecture courses.

For the first few years, residents gave their opinions on each course as they completed it, either on written forms or in informal interviews and reports which were used by the Curriculum Committee of the School in working out revisions. During the years of this project, the last full-scale evaluation of courses by the residents took place in the winter of 1949-50.\* After that, the curriculum was relatively stabilized, and special evaluations were held only when new faculty members were involved in the teaching of a course. The first-year program dealt mainly with history-taking and closed-ward management and some aspects of administrative psychiatry; the second year, with psychotherapy, group therapy, and other types of individual treatment; and the third with special areas, such as child psychiatry, psychosomatic medicine, and outpatient psychotherapy.

### CLINICAL ASSIGNMENTS

Every resident had at all times a clinical assignment. Here he would spend most of his working day, dividing his time between ward or clinic duties and attendance at lectures, seminars, and didactic conferences. Some of the teaching was carried to him right on the ward by way of clinical supervision, both formal and informal. Although there was necessarily an institutional differentiation between clinical assignment and instruction—indeed, he received a salary for the one and paid tuition (or had it paid for him by VA) for the other—in his working day the lines of demarcation might be much less clear.

\* They were resumed later and are currently being used.

At first, residents were rotated through all psychiatric and neurological services in Winter Hospital, spending three months on each. Since it proved hard on the patients to have residents change so frequently, after June 1948 residents spent six months on an assignment. At the same time, it was decided that residents were to enter the program only in July; each one signed a contract for a full year, which could be broken only by mutual agreement. This change reduced the confusion caused by residents' entering or leaving at various times during the year.

The organizational structure of the hospital changed from time to time, but during most of the period of our research it was set up in terms of four services: Psychiatric, Neurological, General Medical, and Surgical. The first of these was the most complex. It comprised an Acute Intensive Treatment Section of closed wards for actively psychotic patients; another Acute Intensive Treatment Section of open wards for neurotic and other less severe conditions; two Continued Treatment Sections for more chronic types of patients, who might have open-ward privileges if psychiatrically indicated; a Women's Section, with both open and closed wards; a Psychosomatic Section, making a bridge between the medical and the psychiatric services, to which was attached a Diagnostic Appraisal Unit for intake evaluations of patients; and—forming a bridge to Neurology—an infirmary for neuropsychiatric patients who were somatically ill or enfeebled by age. The Neurological Service also had a Neurosurgical Section.

From the beginning, a few second- and third-year residents had opportunities to take part in the study of outpatients at the Menninger Clinic, and a few others to work in the Menninger Hospital. At most, two residents at a time could be accommodated at the Southard School for disturbed children, where they participated primarily in evaluations of outpatients, also spending some time at the state Boys' Industrial School for delinquents in North Topeka. Men who were assigned to the Menninger Clinic or the Southard School usually spent part time in the City-County Clinics and later in the Shawnee County Guidance Clinic, where low-cost psychiatric services were offered to the community.

A VA Mental Hygiene Clinic was organized in Topeka at the beginning of 1948; a little later on, its staff also had charge of a similar clinic in Kansas City. These clinics (staffed entirely by graduates of the School) provided clinical placements for third-year resi-

dents, which were highly prized by those who were especially interested in learning to do psychotherapy with outpatients.

A similar type of experience was made possible through the Student Health Service of Kansas University, in the nearby city of Lawrence (thirty miles away). This was also staffed by men trained in Topeka. Later, assignments became available for residents from Winter in the Children's Department and in the Outpatient Department of the Topeka State Hospital.

Beginning in mid-1950, there were ten training residencies for Fellows of the Menninger School of Psychiatry in the Topeka State Hospital, which had just been developed into an accredited teaching institution. Within a very few years this hospital was transformed from what was little more than a custodial institution for mental cases into one of the top-ranking state hospitals of the country. All of the professional staffs increased greatly in numbers and in qualifications. The atmosphere of the place changed from one of hopeless and oppressive quiet, where patients sat all day in rocking chairs and simply rocked, to one of relative cheerfulness, hope, and constructive activity. Menninger Foundation personnel participated in the program of improvement and in some degree Topeka State Hospital replaced Winter, which was by then an established going concern, as the frontier—the locus of fresh excitement and enthusiasm. The Fellows who did their clinical work there attended the same courses and seminars as the other Menninger School of Psychiatry Fellows and therefore all were able to share their experiences in discussions.

#### METHODS OF TEACHING

In describing some of the various methods of teaching that came into use during the period of our study, it is less our purpose to catalogue the School's program than to add a few strokes to the picture of the residents' working environment that we hope will emerge from the first section of this book. By the time our final groups of subjects were being trained (1949-53), the following various teaching methods were being used, in addition to the lecture sessions. Perhaps a quarter of these methods were already in use by the end of 1946. The rest—and others which were tried and discarded—developed with the growth of experience in large-scale psychiatric training.

#### CLINICAL SUPERVISION—INDIVIDUAL

By the middle of 1948, all services at Winter had been organized for teaching, each being staffed by a qualified chief and one or more assistants, with at least one senior social worker and at least one senior clinical psychologist; residents on each service were supervised in part by a group of attending physicians on regular schedule. A resident usually had regular appointments several times a week with a service chief or assistant chief. This psychiatric staff member would advise and guide him in his psychotherapeutic work with a particular patient, help him to schedule and delegate his paper work, discuss with him means of setting up a therapeutically valuable ward organization managed by the patients themselves, and go over many other work problems. Attending physicians from The Menninger Foundation would come to a service for a whole morning several times a week, dividing this time among its wards. On each ward, one or more residents would meet with the attending physician, generally presenting a concrete psychiatric problem to him. The consultant listened, asked questions for more information, and sometimes interviewed the patients. Other staff members on the psychiatric team were often present, and together the small group worked out the diagnosis, the treatment program, the strategy of management, or whatever was required. On some services the consultant or attending physician provided individual supervision of psychotherapy. Usually there was more than one consultant to a service, but staff members still had to carry the main burden of individual supervision.

Since each resident usually had from a dozen to two dozen patients under his care, he had to use his own judgment and try to apply the general principles he was being taught. Usually, however, there would be two residents on a ward, often with a psychologist or social worker, and always with nurses and aides. Depending on the ability and willingness of the resident to become a student again, he could learn a great deal from working with these colleagues, many of whom had had years of experience in the hospital treatment of patients.

The residents attributed much of their learning to individual supervision and usually wanted more of it even though considerable training time was already devoted to it.

#### CASE CONFERENCES

At a teaching conference, a resident would present the abstract of a case, usually for assistance in diagnosis or treatment planning. A consultant, attending physician or hospital staff psychiatrist directed the meeting, and usually all members of the psychiatric team participated along with other residents. After the patient had been brought in for a brief interview, there was general discussion. Sometimes (particularly at the beginning), such conferences were attended by all residents; later, smaller case conferences were held on each of the teaching services of the hospital. With decentralization came also a good deal of diversity: sometimes residents were required to read prepared papers, at other times to speak extemporaneously; some gatherings were set up specifically as team meetings, whereas at others only psychiatrists and residents attended.

#### DEMONSTRATIONS OF INTERVIEWING AND PSYCHOTHERAPY

In 1947, after construction of a large one-way vision room with provisions for fifty or more persons to watch and listen to what went on in the main room, two new types of teaching were initiated. Two faculty members began seeing patients for weekly therapeutic appointments in this room where residents could observe and—after the patient was dismissed—discuss with the therapist anything that he or they wanted to bring up about the treatment. And in an interviewing seminar, faculty members would begin by interviewing patients; in subsequent sessions, the residents would take turns conducting intake and diagnostic interviews with a variety of patients, the rest of the group looking on. The usual pattern for these meetings was a relatively short interview, followed by group discussion.

This teaching method was among the most popular with the residents.

#### ADMINISTRATIVE CONFERENCES

Although “administrative conference” may not sound like a teaching method, attending such meetings on the sections was an important educational experience for residents. In them they learned what kinds of administrative problems came up and how they were handled in a hospital setting like Winter. On some services, the chiefs or assistant chiefs found it possible to get through their agenda with such dispatch that a good deal of the time in these meetings could

be devoted to more general discussions of administrative problems and philosophies.

#### GROUP CONTROLS

This method was used from the beginning primarily as the main-stay of the psychotherapy teaching program. In a typical control—*i.e.*, supervision—group, a faculty member would meet with about a dozen residents, one of whom would present each week an account of his psychotherapeutic interviews with a particular patient. A presenter would usually continue over several months, so that members of the group became quite familiar with the case and with the presenter's method of treatment. These groups almost always were characterized by informality, free discussions, and of critical comment directed toward the presenter. They were fun to attend as well as instructive, and the residents usually urged that their use be extended.

#### READING AND DISCUSSION GROUPS

Some courses that had originally been taught as large lecture sessions were later broken up into smaller preceptorial groups in which there could be more active discussion of the assigned readings under the leadership of faculty men. In their reports and discussions, the residents learned from each other as well as from the designated teacher.

#### SPECIAL GROUP MEETINGS

The teaching programs on all services were generally supervised and co-ordinated by the Director of Professional Education, flexibly enough so that experimentation with different types of instructional meetings was encouraged. At times, for example, consultants to a service were asked to devote some of their visits to conducting seminars on topics in which they were especially competent; some services held diagnostic and appraisal conferences in which the skill of arriving at a working diagnosis and an immediate plan for the patient was emphasized and taught; still others had special conferences on milieu therapy. Many of the better innovations survived; the others quietly disappeared.

#### FORUMS

Every week, during most of 1946, and less frequently but still at least once a month in subsequent years, there were general forums

## 30 *The Setting and the Subjects*

at which the speaker was usually someone from outside Topeka. The forum speakers for the year 1949-50 were probably fairly representative. They included, besides eight psychiatrists and two psychoanalysts, two research psychologists, a social philosopher, two missionaries, three neurophysiologists, an anatomist, a social worker, a dentist, a sociologist, an internist, a speech pathologist, and a *sui generis*—Helen Keller.

### REVIEW SESSIONS

Staff neurologists from The Menninger Foundation or Winter Hospital, or both, almost always offered a series of extracurricular but highly valued review courses in neuroanatomy, neuropathology, and clinical neurology for Fellows and staff psychiatrists who were preparing to take their Board examinations.

### SPECIAL EVENTS AND MISCELLANEOUS TEACHING DEVICES

Attempts were made in a variety of ways to capitalize on the presence in Topeka of visiting experts of one kind or another. An enterprising service chief, upon hearing that a well-known exponent of a school of psychotherapy not taught in the School was coming to town, would make arrangements to have this visitor come to his service and give an informal talk to the residents. Or when a psychiatrist from the Army happened to be in Topeka, arrangements might be made for a meeting in which all residents who wished could participate. For as long as he was in Topeka, Dr. David Rapaport had a continuous course on diagnostic psychological testing, meeting every week. This was elective and varied a good deal in content from time to time; it was focused mainly on giving the theoretical rationale and background of diagnostic testing, but with many digressions on scientific and other issues. In 1946, Dr. Karl Menninger began having an even more informally organized series of talks with the resident body as a whole on Saturday mornings. This came to be known as "Dr. Karl's Colloquium." Here he would introduce visitors and have them talk with the residents, or discuss books that he had been reading, or talk about anything that interested him. The "colloquium" was a highly individual method of teaching, the essence of which only a more anecdotal account than ours could communicate.

Scientific films were shown from time to time at Winter for the residents and professional staff, many of them psychiatric, others documentaries of various kinds—the term "scientific" was given a

rather broad interpretation. Instruction in scientific writing was available a good deal of the time in the form of individual consultations with an editorial worker, who would go over papers written by each resident on assignment as practice writing and discuss with him his strong and weak points in this kind of communication. The teaching of others (younger residents, nurses) is frequently a required part of the duties and the learning experience of senior (third-year) residents in other programs. It never played a large role in the Menninger School of Psychiatry, though teaching opportunities in the training programs of allied disciplines were always available for residents who wanted to take advantage of them. In addition, the School maintained a list of volunteer speakers, including residents as well as staff, which was drawn on to help satisfy the constant requests from groups in the community for speakers on psychiatric topics.

#### VOLUNTARY SEMINARS

The voluntary seminars began very early, when staff members of The Menninger Foundation determined to preserve some of the values of intimate and informal discussions that the forums had once had before the educational programs burgeoned to such size that these meetings necessarily became rather formal public occasions. With staff encouragement, trainees and staff members formed interdisciplinary groups in January 1947 on the basis of common interests, choosing their own topics and inviting senior members of the local staffs to lead them. The topics of the seminars in which residents played a leading part were: development and function of the normal ego; child psychology; inter-relationship of psychiatry and psychiatric social work; supportive psychotherapy; interpretation of myths; the psychology of women; hypnotherapy; and discussion and criticism of various schools of psychodynamics and psychotherapy. Reports on reading by participants made up the core of the sessions. In April, a second group of seminars on similarly diverse topics was set up in the same way, with considerable re-shuffling of group memberships.

So ambitious a schedule took up too much time to be continued on this scale; for faculty members had been devoting more time to the voluntary seminars than to the regular curriculum. Nevertheless, certain residents were so determined to pursue topics that fell outside the regular curriculum that they organized and continued five seminars for which they managed to get faculty leaders. These groups were concerned with group dynamics and social psychiatry,

## *32 The Setting and the Subjects*

lobotomy research, clinical pastoral treatment, scientific methodology, and psychological aspects of imaginative literature. As the curriculum of the School developed, many of the topics around which the early seminars were organized were incorporated into the formal program. But voluntary seminars continued to be formed throughout the period of our study, on psychiatry and religion, on prose or poetic literature, and on some aspect of philosophy or methodology.

## THE ENTERING RESIDENT

WHAT SORT OF MEN WERE WE DEALING WITH—WHAT WERE OUR SUBJECTS like? We shall describe this group of physicians in terms of their origins, motives, and goals, bringing in some test findings and a variety of other data. No one type is delineated, but the group was homogeneous in a number of ways. These common features are undoubtedly due in part to the selection policies of medical schools, in part to forces of self-selection such as the applicants' conceptions of what various kinds of medical practice are like and what different training programs demand and offer.

We do have some evidence about the conceptions of the Menninger School of Psychiatry held by a number of applicants, which throws some light on the kinds of persons who were attracted and repelled by it. The aspects of the School that exerted perhaps the greatest attraction were its reputation as a psychoanalytically oriented training center and its policy of offering intensive academic training along with, and integrated into, the clinical experience offered. To some who were seeking psychiatric residencies these very characteristics were grounds for excluding the School from consideration. In answer to a questionnaire sent in 1953 by Dr. Bernard Hall to a sample of candidates who had begun but never completed application, some reported that they were deterred because analytic

training appeared to them too difficult to obtain in Topeka. Others believed the program to be too one-sidedly psychoanalytical in orientation or insufficiently "organic" and eclectic. The largest number, however, checked as their reason for not pursuing their applications, "the expense of travel to Topeka to complete application procedure." (After this, the Admissions Committee permitted some applicants to be assessed in centers closer to where they lived.) Other deterrents were the seeming formidability of the School's selection procedures, which required two days in attendance with many hours of interviews; reluctance to live in Topeka; the feeling that the size of the institution made for too little supervision; and insufficient income. These reasons suggest that the requirements may have screened out some applicants without strong motivation. Over half of those who replied had become residents in psychiatry elsewhere; of the remainder, the largest numbers had become residents in other specialties or had gone into general practice or military service.

#### GENERAL NATURE OF THE RESEARCH GROUP

The Menninger School of Psychiatry was willing to consider for admission any applicant who had completed one year of medical internship, regardless of sex, color, creed, or national origin. Its criterion was solely that, in the opinion of its Admissions Committee, the applicant should have potentialities for development as a psychiatrist and the ability to withstand the stress of training. For a time it appeared that the Veterans Administration might require all residents in its hospitals to be war veterans, but such a ruling was never actually adopted. Because all VA staff came under Civil Service regulations, residents at Winter Hospital had to be citizens of the United States, by birth or naturalization. It was possible, however, for a non-citizen to work there as a "guest resident" without pay. In addition, the School always had a few special clinical assignments—and, after the opening of the Topeka State Hospital training program in 1950, an increasing number—that were not subject to the citizenship requirement. Therefore, during the period covered by our research, the roster of Menninger Fellows included not only North and South American men and women but also a few nationals of Japan, China, Turkey, Egypt, and other distant places, as well as of several European countries.

The group we adopted as subjects for our research was more re-

stricted, however. We eliminated women and Negroes because, as members of small minority groups within the medical profession, they might have presented special research difficulties, and they were too few to enable us to study these problems. We decided that only those applicants about whom the Admissions Committee had reached a final decision by December 31, 1949, were to be included in our research group. As very few residents of Topeka State Hospital fell within this time limitation, and as differences between the two hospitals might have affected research findings in ways we could not measure with so few cases, we decided to eliminate them and also the few in special clinical assignments outside either hospital. As a result, our study applies only to white, male Fellows who were accepted for Winter Hospital residencies during the first five years of the School's existence.\* This group constituted 86 per cent of all the Fellows who entered the School during the period covered. For purposes of comparison, we compiled data also on applicants rejected by the School, the sources of these data being not only their application materials but also questionnaires sent to them periodically.

The application procedures and the methods by which the Committee decided whether to accept or reject an applicant are described in Chap. 9. From the time the School opened until the end of 1949, when our research population was complete, 456 candidates in the categories covered by our research went through the School's admittance procedures, and 284—62 per cent—were accepted. Of the 284 accepted, 46 withdrew without entering, in most cases in order to accept residencies elsewhere. Thus our research population included 238 Topeka residents, distributed over six classes as shown in Table I-4.1.†

#### INTER-CLASS DIFFERENCES

We have already pointed out in Chap. 3 some of the differences between the early and the later classes of residents. The men in the

\* We also found that residents who came from Latin America had such different cultural backgrounds that it was difficult to evaluate them in the same way, and they too were eliminated.

† Figures published elsewhere (*e.g.*, Winslow, 1946) seem to show hundreds more applicants than are listed in Table I-4.1, especially in the early years. The discrepancy arises from the fact that there were several written inquiries for every man who presented himself in Topeka to be assessed. Our research population includes only those tested and interviewed in person.

TABLE I-4.1. DISTRIBUTION OF RESEARCH SUBJECTS BY CLASSES

CLASS	DATE OF ENTRY	NUMBER OF APPLICANTS	ACCEPTED, ENTERED	ACCEPTED BUT DID NOT	
				ENTER	REJECTED
I	Jan., April 1946	85	49	6	27
II	July, Oct. 1946	101	56	1	44
III	July 1947	66	37	6	23
IV	July 1948	79	32	11	36
V	July 1949	93	46	20	27
VI	July 1950 <sup>a</sup>	35	18	2	15
	TOTALS	456	238	46	172

<sup>a</sup> Only those applicants who were assessed by December 31, 1949 were included in the study; therefore, all of Class VI is not represented.

first three classes had almost all completed their military obligations when they came to the School and often had a year of residency credit to show for it. (As a result, their average length of stay in the School was shorter; see Table 4.2 in Appendix 4.1.) As a group, they were more experienced physicians than the residents who entered later. Many of them had had their introduction to dynamic psychiatry during the war and, upon demobilization, had applied to places all over the country for specialized training. They were of course older, on the average, than the men in the later classes, most of whom came to the School directly from internships. The men in Class I averaged over thirty-three years of age at the time of admission to the School, whereas those in Class VI averaged twenty-eight. Starting with the 1948 class, more men entered knowing that they might be called to render two years of military service before completing residency. For some this prospect became a reality. In 1950 and thereafter, when the law drafting physicians went into effect, entering residents and those already in the School were generally under this uncertainty. (These influences were also largely responsible for the increased proportions of Class IV and V applicants who were accepted but withdrew their applications.)

In each of the first four classes, over 80 per cent were married at the time of application; in Class V, only about 60 per cent were married; in Class VI, fewer still. Most of the marriages of even the older men were recent at the time of application; therefore, although to some extent the difference in the proportions of married residents reflects their age difference, it reflects also a national situation resulting from the disruptions of war. (The national marriage rate reached

an all-time high in 1946, when men released from military service were free to marry, and receded rapidly thereafter.)

#### EARLY BACKGROUND \*

The residents who were our research subjects came predominantly of urban, middle-class families, most of them from the larger cities of the Middlewest or the Middle Atlantic states. Family incomes, according to the residents' recollections, appear in most cases to have been quite stable throughout adolescence and, although not large, still well above average. The precise income figures cited by the residents may be of limited value, as they are based on recollections of varying reliability and apply to varying periods (although predominantly the 1930's), but their range is of interest and probably gives a not-too-inaccurate general impression of the places of their families in the economic scale. Only 8 per cent recalled their fathers' incomes during their own adolescence as being under \$2000 (in the middle thirties over 80 per cent of the nation's families had incomes below \$2000, compared with 30 to 40 per cent at the time we were questioning residents). About half recalled incomes of between \$2000 and \$7500; about a third, incomes of \$7000 to \$15,000; and 7 per cent, incomes above \$15,000.

The fathers' occupations were quite diverse, but large proportions were professional (30 per cent) or managerial (29 per cent), and only 12 per cent were "working men"—skilled or semiskilled workers. Of those in managerial occupations, only a few—2 per cent of all the fathers—could be classified under "big business"; over half the managerial group were retail store owners. The 30 per cent of fathers in professions included 13 per cent who were physicians, 2 per cent being psychiatrists. Worthy of note, although already indicated by the size of the professional and managerial groups, is the fact that 52 per cent of the fathers were self-employed, a very much higher percentage than would be found in a representative national sample of employed males.

\* Many of the data in the rest of this chapter were obtained by means of a questionnaire ("Supplementary Face Sheet") adopted after our study was well under way in an attempt to obtain more personal history than was systematically obtained in the School's application forms. All who applied to the School after the new questionnaire was adopted, and all then in residence, were asked to fill it out. A total of 235, of whom 185 were or became residents, did so. The 185 appear to be adequately representative of the 238 residents in the study.

The record of the families' moves from one city to another is of some interest as a measure of stability. Almost two-thirds of the residents had lived at least the first nine years of their lives in the towns or cities where they were born; about half of the remainder had moved only once during that age period; a very few had moved more than twice. Somewhat more of them had experienced transplantation in adolescence—about a quarter of them once, about a fifth more than once.

Eighty-five per cent reported that their parents had never been divorced or maritally separated, 5 per cent that one parent had died (and in every case that the other had remarried), 10 per cent that their parents had been divorced or separated.

Among the married subjects themselves, none reported divorces or separations.\* Most of them had chosen wives whose backgrounds, so far as they were indicated by fathers' occupations and the wives' educational attainments, were similar to their own. The wives of almost half the married residents had been in professional occupations before marriage—chiefly nursing or social work.

The picture that emerges from these various data is a rather conventional one of a socially and economically stable, middle-class background. It takes on more color, perhaps, when we add to it the fact that, although the great majority of the residents were American-born, 44 per cent had parents one or both of whom were foreign-born. This means that the proportion of first-generation Americans was decidedly greater among our residents than in the population as a whole. The distribution of religious origins was also markedly different. About half came from Protestant homes, a ninth from Catholic, and a third from Jewish (of the population at large 20 per cent are Catholic, 3 per cent Jewish). How this compares with the religious distribution of medical students generally at that time we do not know. We find on comparing our research classes with four more recent classes at the School that, although these proportions remained quite constant through the research classes, in the later classes the proportion of Catholics has risen slightly and that of Jews has dropped sharply (from 35 down to 14 per cent). We have no evidence to indicate what may have caused this shift. Statistically, a change of this magnitude could be accounted for by chance.

\* A few divorced residents are known to us from among those who did not fill out the questionnaire.

Such fluctuations suggest caution in interpreting these figures, yet it is tempting to try to account by conjecture for the high proportions of first-generation Americans and of Jews in the research group. Has psychiatry a particular attraction for people with a special vantage point astride two cultures? Has it acquired so much more respectability as a medical specialty than it had during the first few postwar years that it now attracts a more representative or conventional assortment of medical school graduates? Was there a concentration of minority-group representation at the Menninger School of Psychiatry because exclusion policies made residencies elsewhere unavailable to such groups, and has the concentration changed at the School because there are now more good residencies elsewhere without ethnic restrictions? Our sample is too small and our information about the population of medical schools and psychiatric residencies too limited to lend substance to such speculations. The later changes in the religious composition of the residents at the School is itself a warning against considering our group of subjects as typical of psychiatric residents generally. In fact, no national data are known to us about the religious or ethnic composition of any medical specialty.

The amount of contact with mental illness experienced by the group during early life is of interest as a possible factor in turning them toward psychiatry. Here again, our information suffers from our inability to make comparisons. Thirty-seven per cent had had some contact with mental illness, most of them throughout childhood and adolescence. In half these cases this contact was reported as being close. Unfortunately, we do not know whether this is an unusually high percentage as compared with the general population or with the medical school population. Since mental illness is very widespread, the chances are that the frequency of contact with it in our residents is not unusually high. Moreover, people going into psychiatry are more likely not only to recognize mental illness in their families and friends but also to acknowledge it. Probably the mere fact of exposure is less important than the meaning it has as a result of pre-existing personal dispositions.

Something similar may be said of experience with serious or prolonged physical illness, which in some kinds of people may enhance a reflective bent. Three of the residents in the research groups had had serious cases of tuberculosis. One may speculate that long periods of introversion brought by an illness of this character might

further an interest in treating the mentally ill. It has been reported (Gildea and Gildea, 1944-45), that three of the great American psychotherapists—Mitchell, Salmon, and Riggs—had each had a long period of forced inactivity due to illness—also, as it happens, tuberculosis.

#### CERTAIN TEST SCORES COMPARED WITH THOSE OF OTHER PHYSICIANS

When they entered the School, our residents had, of course, already surmounted the long and arduous series of educational hurdles that we spoke of earlier. It was to be expected that they would score high on intelligence tests, and as a group they did. Their average IQ as measured by the Wechsler-Bellevue Scale of Adult Intelligence, which was administered by the School to all applicants, was 128, a score that is equalled or exceeded by only about 2 per cent of the general population and is therefore rated as "very superior." Scores ranged from 110 to 145. IQ proved not to be a decisive factor in gaining admittance to the School; the average of those who were accepted was only slightly higher than of those who were rejected, and the range was very much the same. (For some comparison with intelligence test results of other groups see Appendix 4.2.)

Much of psychiatric treatment requires talking with patients, so we were especially interested in the scores made on the verbal subtests. Our residents scored somewhat higher on these than on the "performance" subtests, the average for the verbal being 131, with a range of 115-145. As Wechsler points out, however, most people at this intellectual level tend to have about the same superiority of verbal over performance IQ's (Wechsler, 1939).

A test on which we found significant differences between our residents and other physicians was the Strong Vocational Interest Blank. In this test the subject rates or checks his preferences among a wide variety of vocations, hobbies, types of people, amusements, and so on. From his answers to the several hundred questions there emerges a pattern of interests which can be compared with the patterns of interests of similarly tested persons who have established themselves in occupations (Strong, 1943).

We found, as we should expect, that in general the interest patterns of our residents resembled those characteristic of men in occupations involving social service and working directly with people,

and were clearly different from those of men in occupations concerned with business and money transactions or conducted in isolation. Specifically, occupations with interest patterns similar to our residents' were those of public administrators, clinical psychologists, and physicians generally. Very unlike our residents' interest patterns were those of sales managers, bankers, purchasing agents, carpenters, farmers.

Drs. Kelly and Fiske, of the Michigan project on selection of clinical psychologists (see Chap. 7), compared the Strong scores of 119 of our residents with those of their subjects and of 63 young physicians just completing internships at the University of Michigan Medical School. They report that the distributions of interest scores of our residents and the clinical psychologists were remarkably similar, and that both were different from that of the group of medical interns.\*

"Six of the Strong scores show this general picture of differences," they write. "As compared with clinical psychologists and [resident] psychiatrists, the interests of young physicians seem to be much more like those of architects, engineers, chemists, production managers, farmers, and aviators . . . [whereas] clinical psychologists and psychiatrists . . . have interests more like those of personnel directors, public administrators, musicians, and advertising men" (Kelly and Fiske, 1951, pp. 30-35). The cultural interests of the psychiatric residents and clinical psychologists may account for their above-average score on the interest patterns of musicians.

We have in our own records a rather special group of doctors with whom we can compare our residents, namely, 43 men who were rejected by the School and who later, as we know from our "follow-up" questionnaires, dropped out of psychiatry. Comparing their Strong scores with our residents' scores, we find significant differences on four keys: The alumni were more like social science teachers and lawyers in their interests than were those who had dropped out; the latter, by comparison, were more like farmers and executives of manufacturing concerns.

It appears, in short, that men who are attracted into medicine may be responding, in part, to rather different aspects of medical

\* Note that this comparison is unlike the ones we made. Kelly and Fiske compared the *distributions* of scores in pairs of groups on the Strong scales, taken one at a time, and the similarities they found refer to the proportions of samples scoring high or low. The comparisons we made involve only *mean* scores on scales.

## 42 *The Setting and the Subjects*

practice, and that psychiatrists are likely to emerge from among those for whom relationship with people is its salient aspect.\*

### EXPECTATIONS REGARDING FUTURE PRACTICE

The same predilections suggested by the Strong scores of our residents appeared again in their preferences (expressed when they applied to the School) for different types of psychiatric practice. Asked to indicate on a checklist the various kinds of psychiatric work they hoped to do on finishing their training, four out of five checked psychotherapy "other than psychoanalysis," and almost as many checked psychoanalysis. Many who checked these subspecialties indicated that they hoped to spend the major part of their time in them (see Table 4.3 in Appendix 4.1). Four out of five checked teaching, although in most cases as a lesser activity. Two-thirds checked psychosomatic medicine, half checked research (in most cases as a minor activity), almost half child psychiatry. Diagnosis and management, neurology, and shock and related therapy were each checked by one-fifth or less, and almost solely as minor activities. Even fewer checked administration or industrial psychiatry.

The main new trend in psychiatry is the growth of private practice, and most of our residents, when they applied for training, were looking forward to private practice (see Table 4.4 in Appendix 4.1). Only 2 per cent said they wanted to go into VA work full time, and only 10 per cent wanted to do it part time (although we shall see in Chap. 6 that quite a number changed their minds after training, at least temporarily).

The remarkable fact about these figures is the striking discrepancy they show between the expressed desires and intentions of the entering residents and the nature of psychiatric practice in the country at large at that time and for the years just preceding. Here is clear evidence of the changing face of psychiatric practice. In a stable profession, we should suppose, the expectations of men entering it would be at least roughly representative of the existing situation. From these data one would never guess that psychoanalysts and psychotherapists were a minority of American psychiatrists in the

\* No inference should be drawn, from this discussion, about the value of the Strong test in predicting success of applicants in psychiatric residencies. This is discussed in Chap. 13. The results of other tests administered to applicants are described in Appendix 4.3.

early forties or that the dominant pattern of practice was administration, diagnosis, management, and somatic therapy, in large institutions. Of course the residents' aims were not very different from the interests prominently represented in the professional work of the School staff. As we shall see later, there is evidence that our subjects were not expressing idle fantasies; they tended after residency to go into just about the kinds of work they showed preference for in our checklist.

At the same time, we asked applicants a series of questions about the incomes they expected to earn as psychiatrists. Two-thirds named figures between \$7000 and \$13,000 as representing the incomes they expected five years after completing their residencies. A fifth expected larger and the rest smaller incomes. Our subjects were entering training during an inflationary period; incomes (and prices) continued to rise subsequently, and these income expectations have to be interpreted in the light of later economic developments. Even taking that fact into account, we may say that our residents appeared to expect a fairly good but not immoderate return on their investment in education, and that if they erred it was on the side of underestimation. In a later chapter we shall see that for the group as a whole these expectations were being met or exceeded within *three* years of finishing training.

Discussion of expected incomes leads us inevitably to consider again their motives for entering psychiatry. Many residents expected to be better off, when they finished training, than their parents, and upward social mobility appeared to be a part of the motivational pattern. Most of the residents before beginning training had already achieved far higher educational levels than their parents, and the pursuit of still further educational attainments involved in most cases considerable sacrifice. It is fair enough to ask how much the present sacrifice was being made for the sake of future monetary gain.

The interviews and tests to which applicants were subjected provided a number of opportunities to uncover evidences of "money-mindedness" in individual cases, and we shall deal with that subject in later chapters. But it is easy to see that, on the whole, income expectations can be at best only a partial explanation for the choice of medicine as a profession, still less so for the choice of psychiatry as the specialty. In 1951, out of twelve medical specialties, psychiatry ranked third from the bottom in terms of net income earned by private practitioners; only internal medicine and pediatrics ranked

lower. Expectation of monetary return is not a satisfactory explanation for the even greater personal and financial outlay most of the residents made, sooner or later, for psychoanalytic training; psychiatrists in private practice earn approximately the same incomes as psychoanalysts.\*

#### ORIGINS OF INTEREST IN PSYCHIATRY

The question of motivation, on which we speculated at some length near the beginning of this chapter, is obviously not a simple one to be dealt with by merely asking applicants, "Why did you choose to go into psychiatry?" Too great an allowance would have to be made, in evaluating the answers, for the fact that some of the chief reasons were partly or completely unconscious, and also for the fact that an applicant would presumably try to put his best foot forward in answering. Later on, several residents told us that only in the course of their own psychoanalyses did they come to realize some of the reasons they chose psychiatry. In Chap. 16 we present some experienced psychiatric educators' views on motivation for psychiatry, especially as it bears on suitability for the work, since they are in a better position than we to make such observations.

The School asked all applicants to write on their application forms "your reasons for wishing training, the origins of your interest in psychiatry, and your future plans in general." This matter-of-fact approach, although it may not produce much information about the applicants' unconscious motivations, has certain uses. The answers provide us with those reasons that residents were aware of and willing to express when they applied. In some cases, they yield interesting facts of the life history and tell us also something about external forces that might direct an individual into psychiatry.

Responses to so general a request cannot very easily be quantified, but several themes recurred with sufficient frequency on the applications to provide us with general impressions. One of these has to do with the role of the medical schools vis-à-vis the influx of doctors into psychiatry, a subject we discussed more generally earlier.

\* This conclusion is derived from a comparison of the income figures reported for psychiatrists in full-time independent private practice by *Medical Economics* (Nov. 1952) with those reported in a survey by Kubie (1950) for psychoanalysts. It seems a fair estimate that in the period 1949-1951 the median gross income from practice of both groups fell in the category \$17,500-\$20,000 a year.

Occasionally an applicant would trace the origins of his interest in psychiatry to his medical training—presumably medical school—as in this answer (from an application dated October, 1947):

I have become interested in interpersonal and social relationships during my medical training. I wish to continue my instruction, and gain an understanding of the individual in the social setting. Psychiatry as the study of the dynamics of the individual is the medical specialty I would prefer to continue in. At my request, my tour of Army service is with the neuropsychiatric service. At present I am taking a three-month orientation course in Army psychiatry, and have applied for a four-month neuropsychiatric course. . . . After separation from the Army . . . if unable to obtain a suitable residency in psychiatry, I will apply for a residency in neuromedicine or internal medicine.

This man saw his interest in psychiatry as an outgrowth of his medical training and had definite plans for pursuing a more traditional kind of practice if psychiatry should prove unavailable to him. More often, however, the growth of interest in psychiatry would be attributed to influences that appeared to lie primarily outside the medical school, whether because they developed later—in the course of general practice, for example, or during a neurological residency—because they had arisen much earlier, in intellectual interests unrelated to medicine, or because they stemmed from chance experiences. In a small group of applications picked at random, we find some of the earliest applicants writing:

Introduction to psychiatry occurred through friendship with an individual quite devoted to it as a hobby. This man injected the subject into discussions with considerable frequency and presented it so intriguingly that interest quite naturally developed. . . .

The origin of my interest in psychiatry lies in my nonmedical interests. Before entering medical school I was for several years engaged in the study of literature and drama. For one year I directed a theater. During another year I was a research assistant in the Department of Economics at — and a graduate student of economics and philosophy. My interest in the social sciences is not an avocation but an integral part of my work and plans. I believe that these interests are closely associated with psychiatry. It seems to me that knowledge of and concern with social problems and determinants would be largely irrelevant to, for example, the proper reduction of a fracture, but a valuable and perhaps indispensable asset to one working with human relationships. I know of no medical specialty other than psychiatry that could fully engage my efforts and interests.

## 46 *The Setting and the Subjects*

I became interested in psychiatry when the psychiatric approach to political science was presented in a course of political science at — University. This interest was enhanced by further reading of psychiatric literature, courses in medical school, and an externship at the — State Hospital. . . .

Another recurrent theme had to do with the applicant's recognition that his medical knowledge had proved inadequate in dealing with many patients, because their illnesses were not primarily physical:

In my senior year in medical school my first contact with outpatient work brought the realization that the multitude of patients suffered from psychogenic disorder for which the formulae and drugs (with which I was glib) were valueless and for whom little was being accomplished. This conviction persisted through the period of internship. I developed a growing dissatisfaction with the painful inadequacy of otherwise well-trained physicians when confronted by patients with emotional and psychic difficulties and their studied avoidance of psychiatric problems.

Still another theme that repeats itself is the development of a desire for self-understanding:

I became interested in the mind and the emotions on trying to analyze my own during my early medical school days. Interest was further warmed by my choice of subjects and library work during the thesis years of medical school. I am a curious fellow; the human mind has always been a fascinating subject and its evolutionary development a topic for my reading. . . . Specific interest and desire for knowledge was aroused by two factors during my military career: (1) [a close relative's] attempt at suicide while I was overseas; (2) the fact that I was a survivor of a ship sunk during the Normandy invasion. There were many opportunities to observe "combat fatigue" cases and latent personality changes brought to light by the experience. . . .

My original interest in psychiatry really stemmed from my attempts to explain to myself the difficulties I encountered in getting through medical school. Attempting to explain these difficulties to myself on some other basis than intellectual inability, I was launched into a period of several years of self-evaluation. This led me into some very interesting conclusions from which I generalized and began to observe my fellow man in a new light. At this time I was unaware that my observations were scratching the surface of the working knowledge of the psychiatrist. My medical school course in psychiatry, which, in general, consisted of a presentation of hopeless, far-advanced psychotics, was of no help in pointing out to me the direction in which

my thoughts and inclinations were leading me. Their true implications escaped me until about six months after I entered the Army. I have Dr. Anderson, then a major in the AAF and instructor in neuropsychiatry at the School of Aviation Medicine . . . to thank for my first full realization that my observations were a part of a psychiatrist's working knowledge, that the practice of psychiatry was not limited to the task of keeping hopeless psychotics vegetating. I was like a child with a new toy; neuropsychiatry appeared to me in a new, exciting light. The first effect of Major Anderson's lectures was to dispel my feelings of inferiority; then came the realization that my difficulties in school were probably the direct result of emotional disturbances the true nature of which I have not yet fathomed, and the introspection in which I had indulged myself had given me enough insight to see me through my medical degree. Then followed the firm conviction that a psychiatrist was worth his salt; that he rendered a real service to mankind. . . .

The last statement quoted is notable for the feeling it conveys of candor and of freedom in expression. For the most part, written statements in an application are likely to be brief and formal and cannot reveal, as well as a fuller account, the way in which external forces and inner motives might operate to bring a man to apply for psychiatric training. We therefore want to quote here a section from a transcribed interview obtained near the end of training (as part of the Small Sample Study; see Chap. 10) when the resident understood that it was clearly for research use and had nothing to do with admittance to or standing in the School. It gives a vivid impression of how a variety of motives, including deep internal ones, intertwining with those based on reality considerations, led to the final choice of psychiatry.

It should be said at once that no single resident can be typical for the group, and we do not wish to mislead the reader into thinking that this one typifies all, or most, or any known part, of the group. We shall point out, following the excerpt, some of the ways in which we know him to be atypical, as well as some of the ways in which he seems fairly representative. We might mention at the outset that this man was one of the small number of residents who began analysis during residency, and that—perhaps in part because of analysis—he was unusually articulate and self-searching.

INTERVIEWER: Suppose you start out by telling me something about what being a psychiatrist means to you, what you get out of it, why you like it.

RESIDENT: I suppose why I like it and what I get out of it and what it means to me can be answered in the same way; it is hard for me to separate those into different categories. The first thing that occurs to me is that it has something to do with warm feeling and liking people. I've always liked to listen to people's difficulties and I get a great feeling of satisfaction if I think I can contribute something to people's happiness just by listening or by offering some suggestions toward that end. That's the—I would say, the core of the thing. Over and beyond that, it's—it does give one a feeling of power—omnipotence is the psychiatric word—but it really isn't the way I conceive it, though I suppose that basically that's what it would be.

INTERVIEWER: Well, what would you say were the advantages of being a psychiatrist?

RESIDENT: I find it hard to answer a question like that. Advantages of being a psychiatrist? You mean, for instance, over being a grocery man, let's say?

INTERVIEWER: All right.

RESIDENT: I said that because I thought immediately of helping an uncle in a grocery store when I was a kid and coming away with a conviction that never in the world would I be a grocer. It's peculiar that I never thought out the advantages of being a psychiatrist. I never weighed the pros and cons of it. I think I was a little too emotionally impelled to become a psychiatrist over anything else; it wasn't too much on an intellectual level where I would contemplate the question "Well, now, shall I become a psychiatrist or shall I do this or do what?"

INTERVIEWER: Could you tell me something of that emotional background of the decision then?

RESIDENT: Yes, I think I can. I was much distressed in the general practice of medicine by my inability to help more than superficially the large number of neurotics I saw. Quite early in my practice I began to realize what the life of a general practitioner with respect to his work was. A great percentage of his patients were going to be neurotics. Fifty per cent of my patients were neurotics to begin with. This jumped to about 90 per cent perhaps because I was a sympathetic listener with lots of patience. And, of course, at that time my attitude might have been influenced by certain neurotic difficulties of my own. These neurotic patients returned and recommended other neurotics seeking some friendly understanding. And I became dissatisfied with myself—with my kind of treatment of these people. It was haphazard and lacked any deep understanding. I realized that to really help them I had to get to know something about what was behind their difficulties. I rebelled against the one-sided organic approach to illness. And I had a violent antipathy to doctors who had a callous approach to patients—like to cadavers. I know something now of what was behind this dislike. I had my own traumatic experiences with doctors in childhood. They were alarmists and spread

their alarm in fertile soil—I mean in my mother who seemed to feed herself on such alarms and she fed me on them. So I probably was in the throes of some kind of childhood neurosis then; was traumatized in a specific kind of way. The stage was set for me to do a lot of self-searching, to try to settle my own difficulties. It probably stimulated a great deal of curiosity in the difficulty of others. Looking over the past, it doesn't surprise me a bit that when I had the opportunity to chuck my general practice during the war I impulsively and eagerly asked for special training in neuropsychiatry.

INTERVIEWER: When was that?

RESIDENT: In 1943. Before that I had done a lot of reading in psychiatric literature. I had read some of Freud's works in the early days of my college training. I recall going through *The Psychopathology of Everyday Life* and a portion of the *Introductory Lectures*. I was curious about one's inner mental workings.

INTERVIEWER: What else had you read in psychiatry?

RESIDENT: Well, I read Menninger's *The Human Mind* in 1932 when it first came out and I was very much impressed with that. It's strange that I find myself here at his school so many years later. Did I say "impressed" with that book? I was thrilled by it! Not long ago I came across notes I jotted down from this book. It was in an intern's handbook I carried about with me. Practical things which I felt would help me in my ward work in dealing with patients.

INTERVIEWER: Were you an intern at the time?

RESIDENT: Let me see. No. I had read *The Human Mind* in my last year at medical school, or perhaps it was when I just graduated and at the start of my internship. Now why didn't I go into psychiatry? It was a curious trick played on me by my readings in counsel or philosophy for doctors, especially Osler's, that turned me from it toward the field of general medicine. Do you know Jack M . . . ?

INTERVIEWER: No, I don't know him.

RESIDENT: Well, he is now quite prominent in the field of psychiatry. He and I attended the same medical school and he was in the class ahead of me. Jack had decided on going into psychiatry. He was the only man in his class, and probably in many classes that preceded and followed him, who was becoming a psychiatrist as soon as he finished the required internship. He tried to persuade me to take a residency in psychiatry with him. At the time I was strongly under the influence of Osler's writings. I recall a passage in one of Osler's papers to the effect that "a doctor should keep his emotions on ice for some years after graduation from school and he should not specialize until he has had ten years of general work in medicine so that he has a broad knowledge of the field." Something like that. I felt Osler was right about this. Ten years of work in medicine before narrowing down to a specialty. I had to know all aspects of the field of medicine. I really worked at it spending summers in hospitals during my medical schooling and learning something about every-

thing—cancer, obstetrics, radium, X-ray, etc. I did general practice when I began to practice and I worked in various kinds of clinics. But I couldn't keep my emotions on ice, got married, had children, and the poverty of my practice in the depression years prevented me from laying aside anything so that I could take time out for a specialty residency. I'd probably still be stuck in that Oslerian rut if the war hadn't come along and given me the opportunity to learn a specialty while receiving some financial support.

INTERVIEWER: Tell me some of the ways in which you think some of your present attitudes and the things you have talked about are related to your early relationships with your parents.

RESIDENT: Well—my feeling for—sympathy for the underdog—the patient, I suppose, is related to my own experiences with doctors and to my mother's frequent allusions to the sufferings of the Jews. She was very moralistic. Not only as far as her ideas went but she put these ideas into action—on herself and on me. One had to do things right for oneself and also to get others to do the right thing. Thus it would be necessary to get some kind of power over others before one could do this. Broken homes came in for a lot of talk and especially a castigation of men who brought this about. As a very small child this was the very thing which happened to us—for about a year—before a reunion was effected. As you might see, there was a strong identification with mother—and her preachings had a profound effect. This very early suffering of mine determined me, I am sure, in the direction of helping people. I guess I didn't do this via religion because of my mother's skepticism which ran deep in her. A need for power, a fear of doctors, a sympathy for people in trouble, a skeptical attitude running through some kind of faith—this sums up, though perhaps not completely, the motive forces behind my choices of medicine and psychiatry. I should add to this a great curiosity.

This resident differs from most of the others in the length of his previous medical experience, which was unusual even for a member of the earliest classes. The period over which he had been interested in psychiatry was also unusually long. Yet in important respects he is fairly representative, especially of the earlier classes. He took a long time to decide on psychiatric training, he knew that few of his friends had chosen psychiatry and that it was not central to the dominant conception of medicine. He recognized that it was not a deliberate conscious weighing of advantages and disadvantages that led him to psychiatry but an interest emerging from self-searching and the influence of friends and events. He expressed some of the motives we believe to be common—for example, a need to help others. And many of his classmates shared the feeling that they might never have gotten out of their own professional ruts if it had not been for the

war and the military's need to make psychiatrists out of any physician with the slightest interest or willingness.

Perhaps the resident is representative in still another important respect. He refers, in the interview, to "certain neurotic difficulties of my own" and suggests that these contributed significantly to his strong desire to understand such difficulties in others. Here, of course, he was speaking with the benefit of hindsight, and we cannot say how most of the entering residents would have evaluated their own mental health. In the next chapter we shall explore (among other things) the mental health of the group as it manifested itself during their training.

# 5

## CHANGES IN THE RESIDENTS DURING PSYCHIATRIC EDUCATION

BY THE END OF TRAINING MOST OF THE RESIDENTS WERE READY FOR independent psychiatric work. We consider in this chapter how many had arrived, in various ways, at the standards set up by the supervisors, what they had learned both in subject matter and attitudes, and how often they received psychiatric treatment along with their training (see also Holt, 1958*b*).

### SOME DEVELOPMENTAL STYLES

In the course of collecting ratings and descriptions of performance periodically throughout the duration of the project, we saw that residents followed different paths in developing psychiatric competence. At first, we distinguished four rough groups simply according to the changes that took place in their ratings, following a rigorous quantitative criterion. In order to discern types reflecting *styles* of developing and learning, we drew also on the supervisors' free descriptions of the residents' work and personalities. Thereby we altered the groups slightly, further subdividing them and strengthening their evidential basis. The main point deserving emphasis about the isolation of these types is that they do not describe static end-

products; they are process measures, based on different tempos and styles of change.\*

In the following list of types, the numbers of residents in the research group classifiable in each category are given in parentheses.†

1. Steady Highs (37)
2. Overestimated (12)
3. Late Developers (29)
4. Journeymen (117)
5. Uneven Workers (9)
6. Incompetents (24)

There was a gratifyingly sizable group of men who had a good deal of ability at the beginning, who were early able to make use of this ability, and who kept up a high level of performance throughout. They typically excelled as psychotherapists, but also typically were good all-round men. We call these *Steady Highs*. They consistently got high ratings, even though throughout their residencies some of them may actually have been operating at a lower level than might have been possible for them because of neurotic inhibitions or insecurity and anxiety. The label *Steady Highs* should not obscure the fact that they did grow tremendously in breadth; it is affixed to them because they started near the top of the scale of competence and stayed there. These Highs were often recognized early by being given more responsibility; as soon as their residencies were completed, jobs in the local professional community were offered them.

Here are a couple of examples of the way *Steady Highs* were evaluated. At the beginning of Dr. Keene's training, supervisors said of him: "He has a remarkable clinical and administrative ability and flair for research activities; he is exceptionally brilliant and a hard worker." Later a supervisor said, "He is very mature, has warmth and emotion which never seem to get out of hand." About Dr. Sayne they said at the first and at the last: "He shows really outstanding psychiatric acumen. He is very young yet, of course; but he is outgoing, always has something to say and it is usually good. No hesitation about asking for help, but he can go off on his own too."

\* For details on how degrees of adequacy were determined, see Chap. 8. Appendix 5.1 presents full particulars about the way these types were constructed. All residents' names in this chapter, as elsewhere in the book, are fictitious, but the quotations are real.

† Ten residents were in the School too briefly to establish their type of development and therefore are omitted from these categories.

A smaller group were initially *Overestimated*; they were described much as were Steady Highs at first, but then were moved downward on the scale. Only rarely did a personal misfortune prevent their continuing their promise; the usual misfortune was that the supervisors took a better look at them. Inspecting the 12 residents whose ratings declined the most, we found they were people who knew how to make a good first impression. They were, in general, older than most of the residents (including the Late Developers), and more experienced; if not actually older, they looked older. At the same time, they were apparently more self-assured, knew how to talk to people, act like "live-wires," and give a strong impression of knowing their way around. But ultimately they failed to live up to their own advertisements; people saw through them. For example, on one of these, the first supervisor said, "He is extremely intelligent and seems capable but tends to try to get through without any effort." Later a supervisor observed that he was depressed and unhappy because psychotherapy seemed so mysterious to him. Another said, "Although superficially warm, there was coldness and cruelty about him. He made a very good initial impression, was able to verbalize well, but then I saw he was keeping up a smooth flow of talk but not revealing much about his work."

If a man at first looked mediocre, but then developed under the stimulation of the School into a person with more psychological mindedness, sensitivity, and independence than had been thought to be latent in him, we called him a *Late Developer* (*cf.* Dr. Alan Gregg's typology presented in his talk "Emergent Ability," 1952).

There is no single point that characterizes everyone in the Late Developer group perfectly (or any of the other types either), except that they improved. Some of the changes in this group had to do with readjustments in supervisors' opinions rather than only with development of the residents' ability. To quote Gregg, "Real ability, like true virtue, exists. And all too frequently for the comfort of all of us, it exists independently, and at times unconcerned as to whether it be recognized or not."

On careful examination, the 29 residents who were rated as having improved the most proved to be mainly people who made a poor first impression. The predominant type was serious, inhibited, needing a long time to overcome shyness and to warm up. They were more than usually schizoid, retiring, and did not speak up in groups easily. They were often very self-effacing people. Perhaps as a group they

knew better than most what they needed, for they got therapy during training more often than other groups: 47 per cent of Late Developers obtained treatment (training analysis or other treatments) as compared with 25 per cent of the Overestimated and 22 per cent of all others. (Of course, this difference may be interpreted the other way: that if a man was treated, he was more likely to improve and *become* a Late Developer.)

One subtype we have labeled *Rough Diamonds*: a little heavy-handed, and blundering at first, they later became quieter as well as more self-expressive. We also had a few pseudo Rough Diamonds who assumed this role as a means of disguising callous, careless, and even sadistic behavior. Such a man never had his rough edges rubbed off in the School, never demonstrated that he possessed previously concealed or germinal qualities of gentleness, sensitivity, and warmth. Many of these people seemed to be putting up a hypermasculine defense and to be too threatened by their own more feminine sides to allow the emergence and development of capacities and traits like empathy and compassion.

Another subtype of the Late Developer is an adult analogue of what we see on the childhood level as a neurotic pseudo imbecile. These men, because of neurotic difficulties that were often deep-seated, looked heavy, dull, and without any spark to begin with, especially if there was a depressive aspect to their neuroses. Yet these same men turned into good, even outstanding, psychiatrists, particularly after treatment.

A somewhat related category consisted of men who were recognized from the beginning to have outstanding ability but were seriously hampered by neurotic difficulties and were unable in their first year or so in the School (or even throughout the period of their residency) to perform at the high level of which they seemed capable, and which they did ultimately more or less achieve—again usually with the help of therapy. Together, these last two groups may be called the *Neurotically Hampered*. (Dr. Abbott, who is discussed at length in Chap. 11, was a Late Developer of this type.)

Self-confidence is the quality that supervisors most frequently said improved; other qualities were mood, clinical ability, maturity, control of emotions, and warmth. A supervisor said about one in the predominant type (early in his stay): "He is personally unprepossessing, . . . sincere, interested, warm feeling for patients . . . some anxiety, but has potentialities." Another supervisor disagreed about

his potentialities: "A good steady average chap who will never set the world on fire." Another resident of this type had this said of him: "Dr. Grove is slow in thought and action; there is so much tact, courtesy, and form in his spoken and written communications that the barrier not only holds up his work but irritates the listener, but he is kindly and patient, and has shown considerable development." Later in residency, these men were described much more glowingly.

The *Journeymen* were the most numerous group; they can be subdivided on the basis of general level of performance into the Middle-of-the-Roaders (76) and the *Plodders* (41). Their level of ability was not high, but they worked steadily with passable competence. Middle-of-the-Roaders comprised most of those who were rated at the mean and near it on both sides. *Plodders* made up a sizable proportion of the second quartile from the bottom and the better half of the poorest quartile.

*Middle-of-the-Roaders* are almost by definition not particularly vivid personalities; supervisors sometimes found it hard to recall them. This group traveled the widest road; for it was the largest, most diverse type. Only one stayed at the School less than a year, and a high percentage went on to bloom *after* residency. A frequent theme of supervisors' statements about this group (and about other types that did as well or better) was that, although they had problems, their work efficiency kept up. Dr. Mode: "Although anxious, he remains effective and intelligent." Sometimes they were more colorful, with a balanced mixture of assets and liabilities. "Dr. Auer is eager, ambitious, mature, older, but still shows some anxiety with patients and is apt to take a grandiose manner."

Middle-of-the-Roaders seem to have a modicum of psychiatric talent, but the *Plodders* may have lacked specific aptitude. They did not change much; often they were quite inconspicuous, not widely known among either the residents or the faculty. They might be regarded with affection but rarely with a great deal of respect. One of the favorite terms applied to them by the initial interviewers was "mediocre." The work of many of them was called incompetent by an occasional supervisor, and six left before completing one year—a larger proportion than in any group except the *Incompetents*.

Many of the *Plodders*, especially those with the poorest performance, were described as having severe emotional problems, the most common symptom being depression. Very few tried for, or got, psy-

chiatric treatment while they were in the School. Only rarely did they quicken their pace after residency. Quite a few were rigid older men; four were rigidly immature. There was a small subgroup who (at assessment) seemed quite sick in a near-psychotic way, but who managed to remain inconspicuous and somehow had an ability to get certain essential aspects of their job done so that they stayed out of unfavorable notice. Because they looked so ordinary and gave a gross appearance of getting along all right, the occasional reports that came in suggesting much more unfavorable evaluation tended to be discounted, overlooked, or simply incorporated into the probably less perceptive but more prevalent "average" rating.

For most of the Plodders, the dull, diligent aspects were stressed in supervisors' evaluations. For example, Dr. Frost: "He is neither brilliant nor dumb. He is an older man who is not very poised but is very conscientious." Dr. Grey "is dull, immature, rigid, uses stereotyped thinking . . . yet meets his responsibilities." Dr. Guest "is depressed, generally inhibited, plodding, dull . . . shows no promise for the future." Dr. Ernst "looks like a little bulldog but doesn't seem to catch on. Yet he's conscientious and earnest."

Most of the staff's original conception of the School seems to have been that it was to be a training place for exceptional men of special talents, who would be primarily leaders in the psychiatric world. Ordinary, adequate, but uninspired and uninspiring practitioners were insufficiently appreciated, and many men who were thought to be of this caliber were turned away. After the first rush of applications was over, it soon became apparent that this conception of the School would have to be changed, and hospital authorities began to be more satisfied to get patient and relatively trouble-free Journeymen who could be counted on to get the basic job done. Dr. Karl Menninger saluted this type in his "Psychiatric Fable" (1948), but he was still expressing a minority view.

Apparently it took Admissions Committees a long time to become reconciled to this point of view, even though the statistics were plain enough: the number of applications was sometimes not very large, so that the needed quotas could not be maintained if many men were rejected. If anything, the somewhat younger groups of physicians who applied as time went on were, because of their very unripeness, less promising-looking than the first applicants had been; yet the term "mediocre" was still used for those to be rejected. Nevertheless, a number of men were taken into the School and were

allowed to graduate in spite of a demonstrated lack of any touch of genius, creativity, or other extraordinary talent.

Then there were the *Uneven Workers*. Some were brilliant men who were recognized from the beginning to have considerable talent, who showed it sporadically, in bursts and flashes, but were never able to maintain a comfortable and reliable level of competence. It seems that the very qualities that enabled them to do extraordinarily good things at some times forced them into puzzlingly egregious behavior at other times.

The Uneven type was often seen differently by different supervisors and by the same supervisor differently at different times. Some were variable without being particularly gifted. No one seemed able to get consistent impressions of them. Many were very immature people who did not come near to finding themselves by the end of the training period. Most of the Unevens were very bright, however, and had talents that inspired respect, along with severe liabilities. For example, Dr. James: "Brilliant but has a flair for antagonizing people. He can't stand supervision. He is impulsive; he makes promises to patients he never intends to keep. A gifted man, but his potentialities remain to be seen." Dr. Hartz: "Flighty, flamboyant . . . has a potentiality for getting into clinical and administrative difficulties . . . bright, intuitive, sensitive yet immature and barges into psychotherapeutic procedures without knowing what he is doing."

It is interesting to speculate on what a different problem these different types of men create for a school, or for a training analyst. They obviously need different relationships to supervisors, different amounts of academic *vs.* on-the-ward teaching, and so forth. In treatment, some of them need to be opened up and to have some of their inhibitions relaxed; others need to be calmed down and to learn control.

When we look at the small minority of poorest residents, we see quite a number of different types of *Incompetents* and near-incompetents. Just as one would expect, they usually left soon after starting in the School, ten of them in less than three months. The first and largest subgroup is the *Very Sick*, people who led unusually disturbed personal lives. Some were severely depressed, and several were very schizoid or near-psychotic personalities. The latter sometimes broke down and were unable to carry on; a few were hospitalized; a few committed suicide. Several underwent long-term psychi-

atric treatment which did not improve their professional work much.

Included among the Incompetents are some with more or less psychopathically irresponsible character disorders, the *Irresponsibles*. There were several residents who might be described diagnostically as neurotic characters who struggled along constantly upsetting their patients and themselves; they were the *Vulnerables*. Miscellaneous types include a few extremely immature men who never seemed to learn what was expected of them, several severely *authoritarian* characters, and some *near-psychotic brilliant* ones who got into various kinds of trouble and whose future seemed completely unpredictable.

Finally, there were a few Incompetents who were simply in the wrong field—the *Uninterested*. They may not have done anything too damaging in their work, but they never quite seemed to get the hang of it and never learned to like it enough to use the possibilities they actually had. Sooner or later they discovered they were really not much interested in psychiatry and took themselves out of it. Sometimes these people seemed to be rather sick under an unimpressive exterior, sometimes they appeared to be just more suited to another specialty.

The supervisors described one of the Very Sick as follows: "He had borderline capacities as a psychiatrist . . . became acutely anxious and depressed following a long-time marital problem. The problems of patients touched him even more painfully than before. He found dealing with psychiatry even in the mildest form almost unbearable to him. In spite of my advice he talked to practically anybody in the hospital who would listen to his problems." Another resident was quite disturbed yet managed to go on in the School making many blunders and espousing odd theories of treatment. He was described by almost everyone as intelligent but eccentric: "He's a very keen and thoughtful man but really very sick; he has a kind of rigidity, for instance about not prescribing sedation for very anxious patients, which reflects to me a brand of schizoid asceticism which no psychiatrist could get away with in general hospital practice. He was extremely conscientious about doing paper work and anxious to do more psychotherapy, more than the time he had available with so many other practical problems in managing patients."

On the basis of the inadequate material we have on applicants who did not enter the Menninger School, or who left it of their own accord before the level of their competence or incompetence could

be recognized, we do not have much way of saying whether all these types might be expected to turn up in any residency program with the same frequency as they did in our research group. No doubt most of them show up sooner or later in any residency training program. What hints we do have from our initial assessment suggest that all types are probably found among rejected candidates, except possibly Steady Highs, who were almost always recognized, or at least were considered acceptable and were offered appointments. Not all of these accepted Fellowships in the Menninger School of Psychiatry, however; potential Steady Highs often had acceptance from several programs, and sometimes they chose to be trained elsewhere. Our impressions are that there must have been many Brilliant Unevens and Late Developers as well as Journeymen among the rejected applicants. Our hope, of course, is that most of the potential Incompetents of all types were weeded out, and it seems likely that the proportion of Plodders in the rejected group is somewhat higher than among alumni.

One interesting and important sidelight on this problem was yielded by a study of subjects who dropped out of psychiatry (most of them after having been rejected by the Menninger School of Psychiatry, though oftentimes they did not give up attempts to go on in the field until a year or more later). We matched these "Drop-outs" against other rejected applicants of similar age, intelligence, and predictive ratings and made a qualitative analysis of the psychological test reports that had been written about members of each group. Between them, these two groups (Drop-outs and Controls) included practically all the applicants who, in the opinion of the Admissions Committee, were quite unsuited to or undesirable for training. It was striking that there were few of the Very Sick in the group of those who did go on in psychiatry; mostly, they looked like Plodders, or at the worst, Irresponsibles and Uninteresteds. Fully half of those who had dropped out of psychiatry appeared to be Very Sick; most of the other Incompetent types were freely represented, judging on the basis of the psychological test report. We know that these test characterizations were far from infallible. Yet a study of test reports on known cases shows that, qualitatively, they were much more accurate than were the quantitative ratings based on them.

Among the applicants, therefore, the ones who had the more serious psychiatric disturbances were apparently unable to obtain psy-

chiatric residencies, or if they got them, to stay with them; even if they did stay with them, they tended not to establish psychiatric practices but to go into internal medicine or general practice. In these fields, their personal quirks are probably less conspicuous than they would have been if they had remained in psychiatry, and also may not be so damaging to their professional work.

#### WHAT THE RESIDENTS LEARNED

So far, we have talked mostly about types of general competence and not so much about the special skills that are involved. The residents usually learned to be diagnosticians, administrators (from taking responsibility in running a ward of patients); they learned how to manage patients' problems and how to deal with relatives. A few of them, who had the opportunity to work at Southard School, developed special skills in child psychiatry. Others made the most of occasional opportunities for learning how to teach. Only a handful were able to get training and experience in research.

The skill that most of the residents worked hardest to perfect was individual psychotherapy. Nevertheless, even the best of them felt at the end of residency that their therapeutic skills had much room for further development. Though relatively little training in group therapy was given, few Fellows complained that this skill was not more emphasized and developed.

One of the richest sources of information about the resident's view of his training is the interview carried out with him by a member of the School's staff when the resident had completed his work and was about to leave Topeka. In these hour-long discussions ("terminal interviews"), each resident was asked for his frank impressions of what the training experience had meant to him—what he had gotten most from it and what he had liked least about the training. The School's administrators found these interviews very useful as a way of keeping in touch with the extent to which they were achieving their aims and satisfying the residents' needs.

The problem of learning psychotherapy was brought up by the residents often in the terminal interviews. Most of the third year and much of the second was devoted to training in various types of psychotherapy, which was taught by a variety of methods (see Chap. 3). Nevertheless, the interest of the residents in becoming proficient psychotherapists was characteristically so strong that they usually

expressed in terminal interviews the wish that they might have had even more of this type of training and recognized that they still had much to learn. Many made suggestions, varying from starting group controls earlier and providing more individual controls to providing a psychoanalysis for everyone.

The residents learned not only skills but also attitudes toward these skills. Such attitudes were not explicitly taught but were gradually absorbed from people already on the scene. Many of them were not local attitudes only, but part of a developing national professional ideology. We find such a hierarchy of values attached to the divers skills in psychiatry in our sociometric data (described in Chap. 8). For example, in 1947 we asked the Fellows to tell us whom they liked among their fellow residents, and who they thought were the best administrators, diagnosticians, and so on. There was so much overlap between their choices of those they considered to be the best psychotherapists, best potential analysts, best friends, and best all-round psychiatrists that this can be thought of as forming the tetrad with the greatest status in their eyes. Those who were chosen as the best administrators were *not* thought to be the best psychiatrists.\*

One would expect that the learning of skills would be accompanied by the acquisition of knowledge, which could be demonstrated by formal examinations. Few such examinations were given in the Menninger School of Psychiatry in its earliest years, but the ones that were given afford us an opportunity to see what kinds of testable knowledge had been imparted. (For a full account of such examinations and their results see Modlin, 1955.)

All residents in the School took an examination in March 1952. The examination was the same for all the residents, regardless of the stage of their training, and was aimed at a first-year level. Its purpose was to test not the students but the curriculum; the Fellows knew this before taking it. The committee that constructed it considered that a resident should have some knowledge of the following subjects by the end of his initial year, but it wanted to know how well all classes could perform in them: history-taking, the psychological ex-

\* For example, the correlation between choices of "best potential psychoanalyst" and "best all-round psychiatrist" was .89; that between "best administrator" and "best all-round psychiatrist" was .15. (Bold type indicates significance at the 1 per cent point; italics indicates significance at the 5 per cent point; when the correlation coefficient is printed in regular type, it is not significantly different from zero. For an explanation of correlation, significance, and the like see Statistical Note, pp. 359-364.)

amination of the patient, elementary psychodynamics and psychopathology, classical psychiatric syndromes, psychiatric diagnosis, history of psychiatry, functions of the psychiatric team and its members, ward management, management of psychiatric emergencies, and the treatment of psychoses, including physical therapies, milieu therapy, and indications for psychotherapy.

Briefly, the results of the examination showed that first-year men excelled in knowledge of matters covered in the first-year curriculum (*e.g.*, questions on the history of psychiatry, or questions on the temperature of a sedative tub or the dosage of current for electro-shock treatment and other aspects of physical therapies). Fellows in the later years of residency did better than their juniors on questions about such matters as interviewing technique and ward management, which played a larger part in their daily professional work than such matters as psychiatric history, etc.

The general trend of the results suggests that a resident picks up many bits of information and skills which stay with him if he needs them in his daily work. Yet even forgotten scraps of knowledge may help modify the background into which they sink, shaping general attitudes, principles of broad applicability, and a deepening clinical judgment. The residents who were nearing the end of training did best when questions were least loaded with specific *facts* and when they required the examinee to function in ways that were comparable to aspects of real practice—*e.g.*, knowledge of ward management, indications for expressive psychotherapy, and criticizing a transcript of specimen interviews.

Thus, these results fall into line with the findings of other studies about the effects of college or other higher education. Long after the former student can no longer recall lines from the *Canterbury Tales*, he tends to retain an appreciation of literature that is the more sensitive for his having once known Chaucer. It is surely more important that a psychiatrist be able to operate with some of the common sense that Adolf Meyer taught than for him to be able to produce on the spot a thumbnail biography of this or any other great figure in psychiatry (see Appendix 5.2).

#### PERSONAL PSYCHIATRIC TREATMENT OF THE RESIDENTS

The various stresses and strains of residency training might be expected to give rise to more serious consequences than dips in

## 64 *The Setting and the Subjects*

morale. Indeed, many residents reacted to the multiple difficulties of their work and personal situations by applying for psychotherapeutic help. How many of the residents sought it, how many and which of them got it, where did they get it? Unfortunately (for scientific purposes) we cannot answer the most interesting question: what did they get out of it?

Almost all of the psychiatric treatment was given through the facilities of The Menninger Foundation, either by the Topeka Institute for Psychoanalysis or The Menninger Foundation's Department of Adult Psychiatry. The few exceptions were some short-term treatments by the Mental Hygiene Clinic at Winter Veterans Hospital and a few nonformalized treatment relationships at WVAH and the Menninger Clinic which did not find their way into any of the records.

Application for treatment was usually left entirely up to the resident, although the resident's personal counselor \* or supervisor (or wife) might encourage him to take the step. The delay before starting any kind of treatment varied a good deal, depending partly on the type of treatment indicated. Supportive treatment was usually available at once, but psychoanalysis always required a wait of from a few months to as high as two years. The applicant for treatment did not usually have much choice from among the available therapists, but the therapist often had a choice among several patients on the waiting list. The resident's preferences, type of problem, financial resources, and the time he expected to remain in Topeka, all might influence the type of treatment and the assignment of the therapist.

For statistical purposes we have divided therapy into two main kinds: "personal" treatment and "training" treatment. The latter refers to didactic or preparatory analyses under the auspices of the Topeka Institute for Psychoanalysis as part of the program of preparation for becoming a psychoanalyst. Those applicants who sought analytic treatment through the Institute were of course judged on the basis of their potential competence as psychoanalysts as well as their suitability for undergoing analysis. Occasionally an applicant was considered too ill to be acceptable as a candidate for training

\* For awhile, there was an advisory system under which each resident was assigned to some member of the Menninger School of Psychiatry faculty who saw him from time to time to advise him on educational (and sometimes, personal) problems.

and was advised to seek a personal analysis and then reapply. "Personal treatment" included in addition to psychoanalysis a variety of psychotherapies, rather difficult to define because of their many subtleties and diversities but sometimes classified as "supportive" and "expressive." They varied from short-term counseling on a specific problem to fairly intensive long-term relationships. The latter might last as long as psychoanalytic treatment (two to four years), although they usually did not (see Appendix 5.3).

Of the 238 residents in the six classes of our experimental group, 22 per cent sought and received psychotherapeutic help of some kind. (This is very similar to the figure in a comparable study of 90 trainees in clinical psychology [Kelly and Fiske, 1951]: 24 per cent of them received treatment.) An additional 43 per cent sought treatment but did not get it during residency. Many of them entered psychotherapy later.\*

Obviously, the desire to advance their professional competence was among the residents' stated reasons for seeking psychiatric treatment. It is a common dictum within the profession that to be a good psychotherapist one must have experienced therapy as a patient; and for those who wish to become full-fledged psychoanalysts, undergoing analysis is a requirement. There was also, in the School, a secondary gain from being under treatment—a gain in prestige and status. We see something similar in certain circles outside the profession; being "in analysis" has a snob appeal which more and more seems to replace the older attitude that there is a stigma attached to needing psychiatric treatment. There the gain in status may come from the current special prestige of the treatment or from displaying that one can afford its high cost. In the School it could come from the fact of being among the few chosen to receive the benefits of the limited facilities available.

Nevertheless, most residents undoubtedly sought treatment primarily because they felt they needed it. Did the great demand signify, then, that our residents were an unusually sick group? There is evidence that we should expect a high rate of need for treatment among the residents simply because they were a sample of physicians. In the words of Dr. W. C. Menninger (1948a):

It may be significant either of the inadequacy of the methods of selection for entrance into medical schools or the philosophy of

\* For further details, see Table 5.1, Appendix 5.3.

being a physician that doctors, taken as a group, are not too well adjusted. In our hospital in Topeka, for instance, with a maximum of fifty-five beds, we have repeatedly been impressed with the high percentage of patients who were either physicians or from the families of physicians. On two occasions, we have had six physicians at one time as hospital patients. Another most vivid evidence of this was the apparent fact that the discharge rate for neuropsychiatric disability was higher for officers in the Medical Corps than in any other corps in the Army.

Beyond their higher morbidity rates as physicians, however, is there any specific predisposition to emotional disorder in psychiatrists as compared to other specialists? Psychiatry attracts people who are in the process of mastering personal problems. It may be from this source principally that one develops an interest in treating people. Eron's findings (1955) on fourth-year medical students at Yale indicate something of the sort: Comparisons among students planning to specialize in psychiatry, pediatrics, internal medicine, and surgery showed that those who chose psychiatry were significantly more anxious than the other three groups.

It would be hasty to conclude on the single basis of how often they sought treatment that our residents were an unusually sick group. Within the group we found that the best were usually the healthiest (see Chap. 16). Most of the residents applied on their own initiative; only a very few were disturbed enough to be hospitalized.

Naturally, psychiatric residents make up a highly psychotherapy-oriented group, so that they seek professional help for problems that many people would consult their friends, relatives, or preachers about—just as it would be natural for a physiotherapist with a minor backache to try massage. Those who get treatment can legitimately think of it as an investment in their professional future, making it possible for them to do better work with their patients and safeguarding their own mental health against the considerable strains imposed by the work. The training atmosphere is a stressful one to many young residents. Working with mental patients forces one to examine oneself constantly and it often becomes a disturbing experience. The resident may become curious to experience a patient's-eye-view of treatment as a way of helping him understand his patients' problems better. Often a resident will overidentify himself with his patients so that he believes he is more like them than he really is. Personality limitations become obvious even in the best of therapists, although

they are more prevalent in the poorer ones. Any limitations in a man's main stock-in-trade as a psychiatrist—his personality—are likely to interfere with his effectiveness, and psychotherapy is the remedy he is usually encouraged to take.

#### CALIBER OF APPLICANTS FOR TREATMENT

Those who applied for personal psychotherapy unsuccessfully were almost all doing poorly as residents; those who got personal psychotherapy were spread among all four quartiles with a slight bunching in the second quartile from the bottom. Those who applied for psychoanalytic training came equally from all quartiles; those who were accepted were rarely in the bottom fourth and most often in the top fourth. (For further details see Appendix 5.3 and Table 5.2.)

It is easy to find a number of plausible reasons why those who did poorly got less treatment, or why those who got less treatment did poorly. We would expect, since psychiatric treatment usually helps people, that it might have made them more effective residents. But why did fewer of the least effective ones get into treatment in the first place? There was a two-way process between the resident and the therapist. Some residents, after discussing beginning treatment with a prospective therapist, decided against going ahead with it. After a man had made one or more such false starts, other therapists may have been inclined to pass over his name in choosing patients for treatment. The fact that the poorest residents were often difficult cases to treat may have been another reason for passing them by; therapists do not want to fail. Also, the lowest-rated Fellows may have had less insight and less recognition of their own need for treatment, and may have been less able to tolerate a process that would develop them. On the other hand, the better residents probably had stronger drives toward health or capacity for personal growth. Finally, the longer the resident or his prospective therapist delayed, the closer came the time of completing residency; and if the resident could find no local job, he had to choose not to start treatment, at least until after he left Topeka. Such jobs were less available for people who performed poorly as residents, and private practice was almost entirely out of the question in Topeka.

Obviously the committee that selected candidates for psychoanalytic training was influenced by the applicant's record of competence, which was routinely requested from the School of Psychiatry. Of course, there may have been immediate benefits from psycho-

analysis that made the treated residents' work better. In addition, being accepted for training may have given added prestige and may have influenced some supervisors' judgments.

We have little concrete information on the *results* of treatment. It seems safe to assume that some of the general growth and maturity we observed was due to the psychotherapy and psychoanalysis a good portion of the resident group experienced. But how can we tell? One apparently simple device would be to identify those men who were treated, ascertain how many of them improved in competence during training, and compare them with the nontreated remainder. The hypothesis would be that those who received psychotherapy would be the most likely to improve the level of their psychiatric performance.

This analysis was made, and its full details are presented in Appendix 5.3. The upshot is easily stated: there was only a slight and insignificant trend in the direction of the hypothesis. Controlling for the length or timing of the treatment produces no substantial improvement in the results.\*

Reflecting on this lack of a finding, we realized that the method made an untenable implicit assumption: that the men who got treatment and those who did not were in other respects comparable. But we have just seen that they were not, and logical considerations would argue that the nontreated group must have been quite heterogeneous. The treated group must often have required the treatment merely to maintain adequate functioning as residents, whereas the untreated group was capable of this without treatment. Moreover, we know clinically that treatment produces at times superficially disruptive results (as in the transference neurosis). With many cases it is an accomplishment merely to keep the patient from breaking down or regressing. Inner personality reconstruction does not necessarily show itself in behavior immediately; it often takes a while to percolate down, as it were, to changes in everyday behavior. In the circum-

\* There is an apparent contradiction in the fact reported that Late Developers received treatment in about twice as many cases as residents who improved less in their performance. But this result comes from examining the data in reverse direction: taking those who changed most and asking how many were treated. Since there were many more who were treated and whose work did *not* improve, we cannot conclude that treatment makes a man bloom late—unless some other (unknown) ingredient is also present. It is more appropriate to a causal analysis to compare improvement in treated and nontreated groups than to approach it the other way. Another possible source of confusion is the just-preceding section, in which it was shown that residents who received psychotherapy were more highly rated than those who did not.

stances, then, we should not expect to find any demonstrable changes in supervisors' ratings.

There is more to say, usually, about sickness than about health, about difficulty than about tranquillity. We have not deliberately neglected the trouble-free, smoothly-developing resident out of any feeling that he makes "poor copy"; but it is hard to give him the prominence that his numbers deserve in a chapter such as this. As a result, some readers may get the impression that psychiatric residents are a remarkably immature or neurosis-prone lot. There is every reason to believe, however, that in these respects they do not differ much from young physicians learning other specialties, or even from people starting work in other learned professions. Indeed, we were impressed by the high degree of healthy maturity in many of our subjects.

Rarely did residents end their stay in the Menninger School of Psychiatry with the feeling of having gained little. One resident said that he had grown more humble during his two years at the School. He explained that he had learned to respect his limitations more, but that the main thing was "a respect for the patient as a person and not just something to treat." Another said: "I came to Topeka with high expectations; most have been realized. In many ways I feel the training was too brief and I know my psychiatric education must continue for many if not all my years." Many, even of those who had done poorly, felt that they had broken new ground. They would agree with Dr. Alan Gregg "that being a good student is not as interesting as forever becoming one more completely."

# 6

## AFTER RESIDENCY

WHEN THE TRAINING YEARS ARE OVER THE RESIDENT IS AT LAST A PSY-chiatrist, and many opportunities are available to him to practice in a variety of ways and settings. We shall not have given an adequate account of the type of men who were our subjects until we have told what they did during the first years after leaving the Menninger School of Psychiatry. Most of our information comes from questionnaires, which were mailed to all subjects annually for the first three years following their departure, or (in case of applicants who did not enter the School) following their assessment in Topeka. (See Appendix 6.1 for a fuller description of the mail follow-up.)

The response to these requests for information was excellent. At least one questionnaire was returned by well over 90 per cent of all subjects, and, though the response fell off somewhat for the second and third years, almost three-fourths of the alumni completed all three questionnaires.

With regard to the rejected applicants, and to those who were accepted but did not enter, our main interest was to learn whether they took training elsewhere or dropped out of the field. It was only among the rejected applicants that a significant number neglected to return even one questionnaire. It is not unlikely, judging from the

experience of other mail surveys, that the proportion who failed to go on in psychiatry would be greater among the nonrespondents than among the respondents; if our findings err, then, it is probably in underestimating the percentage of rejected applicants who left the field.

#### COMPLETION OF RESIDENCY TRAINING AND CERTIFICATION

An earlier table (Table 4.2) shows that 177 out of our 238 residents—74 per cent—completed three years of residency at the Menninger School. Some left to enter residencies elsewhere, some had their training interrupted by the draft, some were dropped by the School, some left for other reasons. At the time of our last reports, however, only 3 per cent had left psychiatry, and these included men who were practicing psychiatry "peripherally"—that is, as an adjunct to general practice or to some other specialty (see Table 6.1 in Appendix 6.1). Six per cent were in training, and the remaining 91 per cent were regularly practicing psychiatry, either as civilians or in the military. A total of 85 per cent had completed the required three years of residency (Table 6.2). If we consider only the first four classes, who would have had more time to complete (sometimes interrupted) training, the figure rises to 94 per cent.

Thus all but a very few of the former residents of the Menninger School of Psychiatry remained in psychiatry. By contrast, among those who were rejected by the School, almost a fourth left the field, and most of these, at last reports, were no longer interested in it. The other three-fourths succeeded in obtaining residencies elsewhere. Only 1 per cent of those who responded to our questionnaires reported that they were practicing psychiatry without residency training.

Regarding certification by the Board we present data for only our first four classes. Most men in the later classes would not have had time to fulfill the Board's prerequisite of two years of full-time psychiatric practice following completion of residency by the time of our final questionnaire. About seven in ten in the first four classes had passed the Board examinations at last report (Table 6.3). Some had delayed taking it and a few were not yet eligible. Of those who took the examinations, about six out of seven passed on the first attempt. Among applicants to these classes rejected by the School, 42 per cent had passed the Boards.

## PSYCHOANALYTIC TRAINING

At last reports, half the alumni of the first four classes had embarked upon the long and costly process of psychoanalytic training—a fourth at the Topeka Institute for Psychoanalysis, a fourth elsewhere (Table 6.4). Many more had applied for training; in fact, only 18 per cent of the four classes reported that they had not applied anywhere. The proportion of applicants showed a decline from one class to the next, however (in Class I all but two men had applied); and as we noted and considered in Chap. 5, applications to the Topeka Institute and other institutes have since fallen off further.

A good many had had or were having therapeutic analyses, and others had undergone other forms of psychotherapy. If we lump all forms of psychotherapy and training analyses together, we find that by the end of our survey four out of every five alumni of Classes I-IV had had treatment (Table 6.5).

## SETTINGS OF PRACTICE

About a fourth of the alumni were working in Topeka at last reports. The rest were fairly well scattered throughout the United States (Table 6.6), with a preponderance in the Midwest, the Middle Atlantic states, and the Pacific states (mainly California). Large and medium-sized cities claimed all but a small proportion, aside from those who remained in Topeka (Table 6.7).

We saw in Chap. 4 that when most of the residents entered training, they planned to go on into private practice. Only 12 per cent planned to give even part time to VA psychiatry after finishing training, and interest in working in state hospitals or other public institutions was small indeed (Table 4.4). Three years after leaving the School, almost 60 per cent of the alumni of the first four classes were doing some private practice; 40 per cent were giving the majority of their time to it (Table 6.8). But one-third were giving time to VA hospitals and clinics (in many cases a major share of their time), and a few were devoting some time to state and other hospitals. Almost 60 per cent were giving time to community clinics, in many cases a preponderant share of their time (the figure is swollen by the number working for The Menninger Foundation, which is here classified as a community clinic).

The Veterans Administration hoped that in supporting residency training in psychiatry it would improve the institutions in which the training was done and, even more, provide trained men to work in its hospitals and clinics. In the first of these objectives the program was obviously successful. About the second, we have seen both overoptimistic and overpessimistic statements.

Our figures imply a more in-between position. In the first year after they had completed training, most of our research subjects were practicing in hospitals, and the proportion in private practice was at least equaled by the proportion working in VA settings. By the second year these ratios had begun to change, and by the third they had changed still more, in favor of private practice. But even after three years the group as a whole was still contributing a substantial share of working time to public institutions, and even if the trend toward private practice should continue, the services rendered during the first few years (and by some beneficiaries of the program probably for much longer) seem to justify the VA policy. This is all the more evident when we compare the record of what our men actually did after leaving training with their attitudes toward institutional practice before they entered training.

#### KINDS OF PRACTICE

The popularity of individual private practice settings would lead us to expect what we found, that alumni were spending considerably more time in practicing psychotherapy than in any other one function (Table 6.9). During the first three years, few spent much time doing psychoanalysis, but the number will rise with the progress of their analytic training (as of the date of our last figures, very few had had time to complete it). A majority of alumni had experience with diagnosis, management, and administration—functions associated mainly with hospital psychiatry—during the first three years of practice, but these were functions they were gradually giving up, and probably the corresponding increase was in the functions of psychotherapy and psychoanalysis.

On the whole, and particularly if we take the trend into account, our young psychiatrists seemed to be doing pretty much the kind of work they had hoped for when they entered training, although, of course, there were exceptions (see Tables 6.9 and 6.10). There was a

good deal of uniformity in the work interests of entering residents; that is, most of them wanted to do psychotherapy, most of them wanted to do some teaching, relatively few were interested in neurology, shock therapy, or the primarily institutional functions. Hence in comparing the work experience of those who desired to do—for example—psychotherapy with those who did not, we are comparing a large group with such a small one that the differences we find can be of little statistical significance. Nevertheless, these differences are rather consistently in the direction of meeting the subjects' wishes—*e.g.*, not many were interested in neurology and not many practiced it, but those who had been interested in it before training were more likely to have worked at it later than those who had not been interested. The major disparities between earlier wishes and later actualities were in the functions of administration, diagnosis, and management, a disparity which, as we have seen, was diminishing with time and increasing private practice, and in the functions of teaching and research. Most of the entering residents had hoped to do some teaching eventually; at last report, slightly more than one-third of those who had said they wanted to teach had had an opportunity to do it. Research was the only major function in which the interests of the entering residents had split in fairly equal proportions; a little over half had expressed a desire to do research, a little less than half had no such wish. Among those who had been interested, almost half reported that they had found an opportunity for it during their first three years of practice. Some, although far fewer, who had not been interested earlier nevertheless did engage in research later on, according to themselves. We qualify these statements because we did not define the term "research" in our questionnaires, and it is susceptible of very flexible interpretation.

In answer to the question, "How well does your present position satisfy your professional needs?" about 4 out of 10 alumni indicated that they were entirely satisfied, 1 out of 10 that he was quite dissatisfied. The others fell somewhere between these extremes. (The dissatisfactions most commonly expressed are summarized in Table 6.11, Appendix 6.1.) The first three have in common a wish to do either more, or a different kind of therapy, and together form the largest category of dissatisfactions. The lack of opportunity to teach is the next largest category. Quotations may show more fully than the tabulation the directions in which the men wanted to move:

If an analytic institute were established here, I would like training, or, if it were possible, to obtain it on a commuting basis in Topeka or Chicago.

Except financially, I feel well rewarded for the present; later I should prefer more psychotherapy and psychoanalysis.

I am quite well satisfied. But I am devoting too much time to practice (for financial reasons) and not enough to teaching. [Another alumnus] and I plan to form a group for better over-all psychiatric care, and for routine interchange of technical and theoretical ideas. But this remains for the future.

[From a man working in hospital psychiatry] It satisfies me only fairly well. If an educational program can be developed, it will be much better. If I can get more staff so that I can be relieved of such work as caring for the physically ill patients and devote more time to psychiatric care, it would be helpful.

#### INCOME

At the time of their last questionnaire, about 1 in 10 subjects of the first four classes was making less than \$7500 a year; about the same proportion more than \$20,000. The median of the reported figures is \$10,700. (The year of report was mostly 1950, 1951, or 1952.)

If we compare the income distribution in this table with the expectations recorded at the beginning of residency regarding income five years after completing training (see Chap. 4), it appears that within the first three postresidency years the group was already doing better than expected (though perhaps not markedly so in view of the inflation in the interim).

It is difficult to judge how the incomes of our young psychiatrists compare with those of psychiatrists generally. Psychiatrists do not make as much money as is commonly believed. The belief probably derives from the psychiatrist's relatively high fee per patient, but it is easy to overlook the small number of his patients. According to the Seventh Medical Economics Survey (November 1952) the median net income of psychiatrists in nonsalaried practice in 1951 was \$13,000, the median gross income \$18,000. The salaried practitioner was reported to earn about \$6000 less a year than the private practitioner. As we saw earlier, in their third year only about 40 per cent of our group were devoting the major part of their working time to private practice. If we adjust the above national medians (gross) accordingly,

the figure with which our group's median income—\$10,700—is to be compared comes to \$14,400. The comparison is, of course, very rough, and in any case we can only conclude from it that probably psychiatrists take a longer time to establish themselves professionally than the period covered by our questionnaires.

#### OUR "TYPES OF DEVELOPERS" IN PRACTICE

In Chap. 5 we characterized our residents according to what we termed the *style* in which they developed during residency—as Uneven Workers, Steady Highs, Late Developers, and so on. Do differences show up among these groups in the later development of their professional careers?

The Uneven Workers tended to have justified the label by trying out more settings and types of work than any of the other groups. The Plodders and Incompetents were most likely to have done shock treatment, the Steady Highs least likely. The Steady Highs had the highest percentage doing psychoanalysis or other psychotherapy. The Plodders were *not* the group most likely to be found doing hospital psychiatry at least part time; the proportions of Late Developers and Steady Highs working in hospital psychiatry were higher. The trend appeared to be instead for the Plodders to go into a type of private practice in which the use of shock treatment was emphasized. If our data were based on larger samples, we might find ourselves concluding that they pointed to a change in the caliber of men attracted into mental hospital work, which not very long ago was reputed to be a haven for men of mediocre talents.

The groups did not differ in satisfaction with their positions, except that the Late Developers were a little more dissatisfied. As regards income, the Overestimated and the Incompetents were somewhat less well rewarded for their work than the others.

P A R T I I

THE RESEARCH:  
ITS METHODS  
AND FINDINGS



## THEORETICAL ISSUES IN SELECTION RESEARCH

AS A SCIENTIFIC STUDY, OUR PROJECT GREW OUT OF TWO TRADITIONS: first, statistical studies in vocational selection and the prediction of success in various kinds of education and training; and second, clinical psychodiagnostics and personality assessment. Our research is the first systematic and large-scale effort to examine and improve the method of selecting physicians for psychiatric training, and in this respect it has been a pioneering work. But it may also be regarded as merely another in a long line of selection researches going back thirty-five years or more. That branch of applied psychology that deals with fitting the right man to the job has a considerable history and a sizable body of accumulated knowledge.

### THE ACTUARIAL TRADITION

"In the first place, no two persons are born exactly alike, but each differs from each in natural endowments, one being suited for one occupation and another for another." This basic insight, enunciated by Plato in *The Republic*, is at least 2200 years old. Yet it was less than half a century ago that the scientific study of "natural endowments" really began, with the work of such pioneers as Francis Galton in England, James McKeen Cattell in this country, and

Alfred Binet in France. The test that Binet produced for the purpose of separating the bright from the dull in the Paris school system quickly proved to have much more general usefulness. It stimulated the growth of the psychological-measurements movement, which had been inspired by Galton, and the prediction of school performance in particular, both of which endeavors were to furnish important insights and research developments for vocational selection. Pioneers in mental measurement worked out statistical method and research design; they tackled the conceptual problem of the nature of abilities and the practical one of measuring them; and all of the work had a direct transfer to the selection of personnel.

On the side of industrial psychology, the fitting of the man to the job began by slow stages with the emergence of scientific management. Such pioneers as Frederick W. Taylor and the Gilbreths showed that the scientific method could be applied to business and industry to replace prevailing practices that had the sanction only of tradition; furthermore, they were able to show that the new approach was financially profitable. Time-and-motion studies and other analyses of what a worker did, and how he could do it better, were followed naturally by attempts to *choose* workers for their ability to carry out the operations. In the earliest years of this century, Hugo Münsterberg was using psychological tests to select such workers as motormen, ship's officers, and telephone operators. A bold program for personnel psychology, including selection, is contained in his 1913 volume, *Psychology and Industrial Efficiency*.

The first World War brought about a surge of new effort in this field. A selection project of mammoth proportions was suddenly thrust on the U. S. Army by this country's entry into the war. A man had to have certain gross abilities and skills to be even a foot soldier, and there were, in addition, scores of specialized jobs in the services requiring special skills. A team of psychologists, experienced in using the individually administered intelligence tests of Binet, set to work to construct an instrument with which they could test the general intelligence of adult men in large numbers. The result was the Army Alpha, the first group test of intelligence. By the end of the war, the Army mental tests had been given to more than 1,750,000 men and officers. Approximately 2 per cent of this total was screened out as unfit for military service.

The special jobs in the Army were filled by having interviewers select men for them. To guide the interviewers, rough job descrip-

tions were made, with physical, educational, intellectual, and "leadership" specifications for each. "Trade tests"—short schedules of orally given questions to test familiarity with the terminology and operations of certain jobs—were also developed for the quick and objective detection of men who were trying to bluff their way into what they saw as "soft spots."

Much of this work was pretty crude, and the tests did not always measure what they were thought to, but the *idea* of determining a man's aptitude for a job before starting him on it took hold and the feasibility of doing so on a tremendous scale was thoroughly established. Personnel psychology was given not only a momentum but an orientation that it has never lost toward dealing with large numbers as efficiently as possible. This has meant, in turn, a preference for objectively scorable tests rather than detailed clinical study of individual cases, which has been the characteristic method of psychiatric diagnosis, the other principal tradition in the assessment of men.

The psychologists in the Army had shown the way, and they were not slow to apply the same methods to personnel problems of business and industry after the war. Many modifications, adaptations, and imitations of the Army Alpha were quickly developed, since different jobs had quite different intellectual requirements. As it developed, not only did a professional person need the highest degree of intelligence, but, conversely, a very high degree of intelligence was undesirable in a routine worker—a file clerk with the same intellectual level as the lawyer for whom she worked was likely to be unhappy in her job. In short, there were optimal ranges for various jobs, with maximum as well as minimum levels of intelligence for best performance.

Industrial psychologists found employment in specific industries and as private consultants, individually and in fairly large-scale collaboration, through such institutes as the Bureau of Salesmanship Research at the Carnegie Institute of Technology and the Life Insurance Sales Research Bureau. They developed many new tests and other methods for the selection and placement of workers in business and industry. Still, efforts were sporadic and scattered; and though some companies began research operations of their own in personnel selection, much of this work was abandoned in the economy cuts which accompanied the great depression of the 1930's.

But the depression itself was instrumental in activating a research project at the University of Minnesota to study general prob-

lems of unemployment. Psychologists of the Minnesota Employment Stabilization Research Institute (established in 1931) studied, among other things, the vocational aptitudes of unemployed persons and the skills and abilities involved in various occupations, and they did some vocational guidance work in addition to developing a number of new tests (Paterson and Darley, 1936). Shortly thereafter, the United States Employment Service set up an Occupational Research Program to facilitate the co-operation of government and industry in similar research on fitting men to jobs. A principal result was the *Dictionary of Occupational Titles*. For the first time, the entire spectrum of occupations was systematically arranged, with detailed descriptions of the actual operations involved in many hundreds of jobs. This was much more of a contribution to selection research than may be evident, since a vital—though often neglected—aspect of any such undertaking is finding out exactly what a job consists of and what is required to do it well.

So far we have talked about the choice of workers by means of intelligence tests and similar instruments. But selection research was rapidly developing other types of tests as well as exploiting additional ways of learning about a prospective worker.

From the beginning of industrial psychology, *performance* tests had been used. In simplest form, they were an attempt to speed up and systematize the old try-out method. Instead of hiring a man and seeing whether he sank or swam, the personnel worker could expose him to a standard situation that contained, in miniature, most of the demands the job itself would make on him. Recent interest in situational tests of the kind used in the OSS assessments (which we describe later) has given many people the impression that they are a new invention; yet as long ago as 1928, M. S. Viteles was helping a power company select electric substation operators by a test (among others) that called for rapid decisions and precise manual adjustments in the midst of the smoking crackle of arcing electrical apparatus—closely simulating an actual breakdown in a substation (Viteles, 1934). At another extreme, potential locksmiths were selected by their performance on a test consisting of a simple lock to be taken apart and put back together.

Another legacy of military psychology to industry was the *personality inventory*. Originally developed by R. S. Woodworth as an aid to psychiatric screening in World War I, the first such device—the Woodworth Personal Data Sheet—was essentially a psychiatric

interview in self-administering form. The subject read each of 116 questions about physical and mental symptoms and checked the answer (Yes or No) that applied to him. The first modification of this device to be widely tried out in vocational selection was the Bernreuter Personality Inventory. Instead of yielding a single score, answers to this test are keyed to give scores on several personality traits, such as self-assurance and introversion-extraversion. Many other such "personality tests" were put on the market, all relying in one way or another on self-report, and mostly with questionable validations, if any; nevertheless, many found enthusiastic proponents among personnel men.

Another widely used type of instrument is the inventory of *vocational interests*. These tests are used mostly in vocational guidance but have been employed in selection, too. The Strong Vocational Interest Blank and the Kuder Preference Test are the best-known species of this genus. E. K. Strong, Jr. reasoned that a person who has a pattern of interests similar to that of lawyers, for instance, should be relatively successful and happy in the legal profession, and similarly for other occupations. His test is like most personality inventories in that the subject checks one of a limited number of possible answers to each question. The answers are scored according to over a dozen keys, each distinguishing as sharply as possible between the interests of one occupational group and those of men in general. Kuder's test is similar from the viewpoint of the subject but is scored in terms of general characteristics of vocational interests (*e.g., persuasive, computational*) rather than patterns of particular occupations.

Despite the rash of new tests developed in the twenties and thirties, business and industry did not by any means unanimously adopt this way of selecting personnel. The *interview* has always been the most widely used technique for sizing up a man for a job and doubtless will continue to be for some time to come. During the first three decades of American personnel psychology, general experience and, to a limited extent, specific research contributed a good deal of useful lore on how to talk to an applicant in such a way as to get a maximum of useful information and impressions with a minimum of subjective bias. We shall discuss the fallacies and pitfalls of interviewing later on, as well as some of its peculiar excellences (see Chap. 17). *Application blanks* or similar informational forms are very widely used and so is the practice of requiring *letters of recommendation*.

*tion*, though the latter are generally thought to be of little value.

Similar methods were developed for choosing among applicants for admission to professional schools, such as those of medicine, law, engineering, and accounting. Almost universally, professional training institutions have selected students on the basis of grades received in preprofessional schools. Where the competition has been stiff for a relatively small number of openings, as in medical schools, an applicant who did not stand fairly high in his college class has had little chance of acceptance. This practice is supported by research findings; grades in premedical school generally correlate well (between .5 and .6) with grades in medical school (Stuit and Schlichter, 1948).

Despite large-scale attempts to devise selection tests, none by itself has been able to take the place of premedical grades as a predictor of performance in medical school. In combination with premedical grades, however, these tests improve predictions (Stuit & Schlichter, 1948; Morris, 1951; Stalnaker, 1950).

The Medical Aptitude Test was developed by psychometricians to supplement grades in screening applicants for medical training and was used in medical schools all over the country. When one thousand medical students of the class of 1932 were divided into two groups—those who had been in the upper three-fourths of their college classes and were also in the upper 75 per cent on the Medical Aptitude Test, as against those who could not get over this dual hurdle—69 per cent of the actual failures fell in the latter group along with only 4 per cent of those who had received high average marks or better in medical school. Later on, however, there was a good deal of variation in the results from year to year and among different medical colleges. In 1947, therefore, the Association of American Medical Colleges recommended the Graduate Record Examination, a general achievement test intended for graduate schools generally. It gave poor results when medical-school grades (except those from freshman year) were used as criteria.

The American Association of Medical Colleges did not give up the attempt to produce an instrument that would test abilities and aptitudes more specifically relevant to performance in medical school. In 1948, the Association asked the Education Testing Service of Princeton, New Jersey to prepare such a test; since that date, the Service has produced, administered, scored, and reported it for the medical colleges. Revisions are made each year, and it is the subject of constant research. The test used in 1950 was taken by 5185 (74 per

cent) of the 7032 students entering medical colleges in the Association. It had four parts, which measured verbal ability, quantitative ability, achievement, and ability to reason in the social and natural sciences, by means of objectively scored questions. Scores range from 200 to 800, with an average of 500.

How well do these scores foreshadow a man's performance in medical school? Since many colleges accept only about 1 applicant in 10, under ordinary conditions of use not enough men with low scores will get in to make it possible to compare their work with that of high-scoring applicants. In 1950, only 29 men who scored below 250 were taken into medical schools and, during the four years, 48 per cent of them failed, dropped out, or did not continue for miscellaneous reasons. Many more (359) who scored between 250 and 350 entered, and a quarter of this group did not graduate on time. By contrast, only about 10 per cent of the thousands who made average scores (500) or better dropped out or fell behind (Stalnaker, 1954). Considering the amount of preselection and the fact that the student group was also screened on the basis of undergraduate grades, recommendations, interviews, and the like, this is impressive evidence that the Medical College Admission Test can make a useful contribution to medical selection.

The status of this and related fields of research just before World War II was summarized in a monograph, *The Prediction of Personal Adjustment*, published by the Social Science Research Council (1941). Paul Horst and his co-authors brought together the methods and findings from work on not only vocational selection but educational achievement, the prediction of marital adjustment, and the success of prisoners released on parole. In this remarkable document they showed the logical similarity of the scientific problems in all these fields and outlined the major steps of method and specific kinds of difficulties that have been encountered in many researches since then, including our own. The following summary draws largely on Horst, supplemented by a few more recent sources.

In the field of industrial selection, the general impression over the years is of ever-expanding successful application of old and new predictive devices to one type of task after another. A recent review of vocational research (Heron, 1954) reports reasonably successful selection of Naval personnel, taxi-cab drivers, machine shop operators, women hosiery workers, and others.

In the related field of educational selection, it had been found

that college achievement could be predicted moderately well by two kinds of scores: secondary school grades and college aptitude tests. A good high school student often made a good college student, just as one might expect; the validity of this predictor averaged about .55 against grades in the freshman year, though it correlated less well with grades in the later years of college. The Scholastic Aptitude Test, which was developed by the College Entrance Examination Board (now the Educational Testing Service), gave varying results at different universities, but it too predicted freshman grades with a validity averaging about .50. A statistical combination of scores on this test with high school grades was somewhat better than either alone, but more than half of the variance in freshman grades remained unaccounted for, and attempts to supplement these predominantly intellectual measures with tests of various other aspects of personality proved fruitless.

Equally good levels of predictive efficiency had been reached by sociologists and psychologists concerned with marital adjustment. Burgess and Cottrell (1939) set up an actuarial table based on findings from a study of over 500 couples; applying it to a check sample of 155 couples, they got a validity coefficient of .48. The data that went into the table consisted of relatively objective background items and data on adjustment during engagement, elicited in interviews and questionnaires. A similar but larger and better-controlled study by Burgess and Wallin (1953), in progress at the time the Horst monograph was written, started with 1000 engaged couples, of whom two-thirds were followed up after three to five years of marriage. A similar actuarial table was constructed and applied to a new sample of 100 men, yielding a validity of .5.

The other predictive studies reviewed by Horst *et al.* came from criminology. During the thirties, a great deal of effort went into attempts to find characteristics of released prisoners that would predict whether or not they would violate parole and relapse into crime. Studies such as those of Burgess and his collaborators made extensive and successful use of the actuarial table technique. To make up such a table, you proceed roughly as follows: First you must collect data on a large number of cases, such facts about each released prisoner as his age, sex, marital status, ethnic origin, type of crime, etc., and also the crucial point: whether or not he violated parole after a given period of time. Next you cross-tabulate each of the personal characteristics against the other until you have a grid of cells, making up all

possible types of prisoners (*e.g.*, young-male-single-Irish-embezzlers; young-female-divorced-Yankee-murderers; old-male-single-Negro-embezzlers; etc.). If you use many characteristics and do not have thousands of cases, some possible combinations will have few or no cases in them. Finally, you compute the success-frequency for the prisoners falling in each cell of the grid, and it becomes an actuarial table. If you know that 35 per cent of 200 young-male-single-Irish-embezzlers have violated parole, the best prediction about another prisoner of that description is that he will *not* violate parole. Such a table is useful only if the proportion of successes differs significantly from cell to cell; it may be necessary to add some kinds of information and discard others until such differences are obtained. Finally, you must try it out on a new group, making predictions which are then checked. If the proportion of success remains about the same as before in each cell, predicting the criterion thereafter is extremely simple; any clerk can do it merely by entering the table with the appropriate description of the case.

The other principal statistical technique of prediction, called multiple regression (or multiple correlation), was most highly developed during World War II. The war furnished a giant impetus to the development of personnel psychology. Again the armed forces had huge problems of screening men for intellectual and emotional fitness, and again a great deal of work went into the perfection of such instruments as the Army General Classification Test and various self-administering tests for the detection of potential neuropsychiatric casualties (Menninger, W. C., 1948a), as well as specific tests for all kinds of specialized jobs.

Perhaps the largest and most productive research effort was concerned with selecting men for the most glamorous and sought-after fighting jobs in the services: piloting airplanes in the Army Air Force. The AAF selection studies were carried out by teams of psychologists, under different leadership, working on selection for a variety of quite different billets in addition to pilot, including radar observers, flexible gunners, and navigators. They have been reported in great detail in a series of fifteen volumes under the general title *AAF Aviation Psychology Program Research Reports*. Only the sketchiest summary of this massive endeavor can be given here.\*

A battery of about twenty tests was chosen (from many more)

\* A longer summary may be found in a series of reviews of these volumes in the *Psychological Bulletin*, 1949, 46, 499-528.

for the selection and classification of aircrew personnel; it included fourteen printed "objective" tests and six tests requiring apparatus. Some of the most successful of the latter were similar in principle to the penny-arcade machines on which one pulls levers and triggers to track down moving representations of "enemy planes." The same tests were generally used, by varying the "weights" or degrees of importance attached to each of them, to select different kinds of personnel. There is some doubt about the specificity of these weights for various kinds of aircrew jobs \* but there can be no doubt that they did the job for which they were intended. Time and again, on successive large samples, weighted scores from the battery of tests would correlate from .5 to .6 pass-fail criteria for the job in question.

In one particularly convincing demonstration, 1,311 aircrew trainees were tested and their scores recorded, but all were allowed to go ahead with flight training to test the predictive efficiency of the battery of tests over the entire range of talent for flying. The result: the composite predictive score, or pilot stanine as it was called, correlated .66 with the criterion of whether the trainees "washed out" or successfully completed flight training. No other attempt to select candidates for training for any kind of job had ever attained such a high level of success with so large a sample of subjects (Du Bois, 1947).

Such was the result achieved using the multiple regression technique. How was it attained? The procedure followed the techniques that had been worked out in many other selection researches, except that samples of a thousand or more subjects could be used, rather than a few dozens or hundreds. Detailed studies were made of what a pilot does, what a student of flying has to be able to learn. Next, psychologists tried to figure out what abilities a person had to have, what temperamental characteristics he should possess, etc., in order to do this job. A large number of tests was assembled, including already standardized measures of some of the abilities in question and a good many new ones. Then large samples of flight trainees were given these tests, and their progress through training was watched. A statistical measure (biserial correlation) was then computed between each test score and the pass-fail criterion. At this point, many tests proved worthless and were discarded. The tests that predicted well were repeatedly cross-validated—tried out on new groups—until

\* In one experiment, each of the weighted combinations for different specialties correlated to about the same degree (from .25 to .40) with the same criterion—pistol target scores on the ground! (Du Bois, 1947)

a stable result had been achieved. Complex statistical methods were used to find the best ways of combining test scores to predict such a criterion as pass-fail in primary flight training as efficiently as possible. Mainly this involved correlating each test with every other and all with the criterion, and from the resulting matrix of figures computing weights for each test score so as to give the best prediction of success or failure. Moreover, the tests were subjected to constant study and revision. With each successive sample of students, the most promising of the originally selected tests were retained, others were modified to try to bring out qualities not yet well measured, new ones were added, and all were again correlated with each other and the criterion to find the most efficient group of tests with the smallest overlap. After this process had been repeated with a few thousand airmen, an efficient battery of reasonably reliable, objective, and easily scorable tests was finally worked out. Thus, a very good level of validity ( $r = .6$ ) was reached and held through repeated trials.

Notice that this statistical procedure has several points in common with the actuarial-table method. It used facts about each subject that were objectively determined—without the exercise of anyone's judgment or clinical skill—from apparatus tests of various mechanical and motor abilities and printed tests of intellectual functions, perception, set and attention, and temperament. It was found that biographical data, collected by printed questionnaire forms and machine-scored, could be used effectively in selecting pilots and navigators; such "personality factors" as were included were measured by this means. After all the tests had been given, a clerical worker could apply the proper statistical formula to anyone's score and mechanically find a good prediction of his probable success.

#### THE CLINICAL TRADITION

In its planning stages the present project was influenced much more by the *clinical* tradition in psychiatry and psychology than by the actuarial tradition. The selection of residents had always been in the hands of psychiatrists themselves, and they naturally brought to this task the techniques and viewpoints they had found useful in their work with patients. Just as education is analogous to therapy (both being organized attempts to modify human beings and their behavior in predetermined ways), choosing a man to train was like choosing a patient to treat. Both processes implied a prediction (a

fact not always clearly seen) that one person would improve or profit from training, another would not. Such predictions were made by psychiatrists and clinical psychologists in general qualitative terms without the use of numbers or precise checking, since there was almost never a clear-cut criterion against which they could be validated.

A critical question is *how* clinical evaluations of personality were made. The psychiatrist relied on the clinical interview primarily in coming to diagnostic and prognostic judgments about patients: talking to the patient, observing his expressions, behavior, and emotional reactions as well as the words in which he told his personal history, the story of his complaints and how they developed. Clinicians differed greatly in the ways they used interviewing and in the kind of information they looked for, depending largely on their own theoretical backgrounds. Nonpsychoanalytic psychiatrists often used directive, probing questions, whereas psychoanalysts were used to a looser, more permissive type of interview, in which the patient's own needs were relied on to guide his associations in a way that would bring out relevant facts.

Regardless of theoretical position, clinicians developed an experience with the far-reaching complexity, intricacy of organization, and subtlety of influences in human personalities and a respect for their tremendous diversity. Since any aspect of a life could show the effects of neurosis or psychosis, the scope of diagnostic inquiry had to be enormously larger than the purview of the personnel psychologist. Having such a wealth of data to deal with, the clinician had to rely on theories with a broad sweep to help him organize and make sense out of all he learned in extended diagnostic and therapeutic contact with patients. The ideas of Freud, particularly, interacted with the diagnostic insights of nonanalytic psychiatrists to create a loose, informal, but rich and useful, system of psychodynamic and structural concepts for the detailed understanding of healthy and pathological personalities. This system was the basic medium of exchange in our own clinical training and that of the other psychologists and psychiatrists who conceived and shaped The Menninger Foundation research on psychiatric selection.

What are the essential elements of this clinical tradition? We have mentioned so far a view of human nature as being protean in its presenting forms, complex in structure, subtle in the interaction of conscious and unconscious determinants, and capable of being

understood only by flexible and detailed inquiry on a broad scale; a conviction that unconscious forces played a tremendous part in determining behavior; a reliance on personality theory, particularly the psychoanalytic, to organize the data; and a faith in unstructured interviewing and qualitative evaluation rather than precise and controlled quantitative measurement. Several other elements should be added. Clinicians are trained to deal with pathology and, accordingly, are more sensitive to malfunction than to efficiency of the person's functioning. Intuitive, nonverbalized judgments are used and respected as often being more useful than anything that can be rationalized; and the focus of interest is on the unique patterning of individual lives, rather than on the properties of particular elements that can be singly abstracted from personalities and studied over many cases. Indeed, the clinician characteristically distrusts statistical studies of isolated functions, which necessarily neglect their embeddedness in concrete personality structures, and sometimes generalizes this attitude to a hostility toward any quantification or statistical treatment of data.\* In the clinical tradition, faith is put mainly in the observer himself as the chief instrument for gathering and processing data.

The clinical tradition originated in European psychiatry and psychology and was applied to nonpsychiatric assessment problems principally in Germany. The best-known of these applications was the work of Simoneit in officer selection for the German army. Simoneit and his co-workers approached the problems with a very different spirit from their American counterparts: where Americans were skeptical and empirical, always demanding to be shown that a procedure had validity and results could be duplicated, the Germans were inclined not to worry about such matters but to emphasize the necessity of getting to know an officer candidate thoroughly in life-like situations, as well as in tests and interviews. They developed a number of ingenious situational tests, which had a good deal of influence on later assessment research.

Many persons played a part in bringing the clinical traditions of Europe to this country, but none was more influential in getting it

\* This last attitude is a misunderstanding, which we do not mean to impute to clinicians generally. It fails to distinguish between the admitted triviality of many studies and the quantitative techniques of verifying propositions that they employ. Statistics can and must be used to test the validity of meaningful predictions and need not produce such triviality (see Meehl, 1954).

into American psychology than Henry A. Murray. He managed to blend the clinical traditions of the psychoanalysts, the assessment techniques of Simoneit, and the contributions of several other theorists, together with much of American pragmatism and experimental method, into a research-based and research-oriented personology, which will be worth our extended attention.

#### THE WORK OF H. A. MURRAY AND HIS COLLABORATORS

In 1938 a book was published which significantly altered contemporary thought about personality and decisively influenced empirical studies in this field. It was called *Explorations in Personality*; on the title page appeared the names of twenty-seven workers at the Harvard Psychological Clinic, in addition to that of Murray, the principal author.\*

In a number of ways the work it reported set a pattern followed by many subsequent researches, including our own. A team of co-workers brought their varied talents and professional skills to bear on a co-operative and intensive study of a relatively small number of subjects. Not only did they pool and share data, attempting to integrate them into rounded pictures of well-understood individual personalities; they agreed to make use of a common theory of personality and to assist in its elaboration and development.

This theory had two striking features: First, it was based upon psychoanalysis but attempted to integrate with it insights from gestalt psychology, Jung's analytical psychology, cultural anthropology, clinical psychology, and the systems of such academic personologists as William McDougall and G. W. Allport. Second, it offered a comprehensive set of concepts for the description of all aspects of personality: motives, abilities, sentiments and values, structural and temperamental traits, dynamically significant aspects of environment, and typical aspects of childhood experience. Much of it was presented in a forbiddingly neologistic terminology, and it had other

\* The tradition of Murray and the Harvard Psychological Clinic affected our research not only through the general impact of these ideas on thinking and research in personality, but in a more personal way. One of the authors participated for two years in a co-operative experimental venture started by Murray at the Harvard Psychological Clinic in 1941, a new set of explorations that had to be abandoned because of the war. As part of this experience, he participated in a group experiment on the diagnosis of personality conducted by Daniel Horn (1943). The design of this unpublished study had a direct influence on the designs of the research in the Office of Strategic Services and the Michigan Project (see Appendix 10.8).

defects; yet this system proved practicable and enormously useful in systematizing and co-ordinating the work of many hands. The concepts were defined operationally enough so that each could be rated quantitatively for every research subject.

At that time, the prevailing tradition in the measurement of personality was the personality questionnaire, or inventory. Usually, more effort went into the choice of items by tedious and precise methods of psychometric item analysis than into finding out what the resulting scores meant. Unimpressed by the practical results of personality-questionnaire research, the workers at the Harvard Psychological Clinic rejected it almost entirely. They invested heavily in projective techniques, which at that time were just coming into use. The new approach emphasized understanding the productions of a particular person in terms of a psychoanalytic theory of forces in the personality. Murray's group devised many such methods, most notably the Thematic Apperception Test (TAT), and subjected them to an intensive clinical development, very different in character from the methods derived from intelligence testing that were applied to the development of personality inventories.

Faith in the ability of the investigators to make their own tools and to understand their results with the help of a theory, rather than relying on statistical refinements and external validation, paid off in impressive insights into individual cases. It is true that the method had the one-sidedness that seems to be the necessary accompaniment of enthusiasm and there was much in what was lightly discarded that might have helped the development of the new instruments. It fell to later workers to supply these refinements, however.

Murray was ready enough to take sides and to utter polemics on these issues. The gage is thrown down at the very beginning of *Explorations in Personality*. Psychologists are divided, Murray said, "into two large classes holding opposite conceptual positions. One group may be called *peripheralists*, the other *centralists*." He goes on to characterize the former group as "objectivists," who are positivistic, mechanistic, or given to physiological explanations; they define personality "in terms of action *qua* action rather than in terms of some central process which the action manifests" and look on personality "as the sum total or product of interacting elements rather than a unity which may, for convenience, be analyzed into parts." Those who enthrone the external stimulus as principal governor of behavior are also included in this group. By contrast, the centralists

(with whom Murray explicitly aligns himself) are concerned with inner organizing, directing, and governing processes in the brain or in the unconscious mind; they are conceptualists, totalists, intuitionists, and dynamicists, believing in the emergence of whole properties, unpredictable from a knowledge of the separate parts, and in abstract, hypothetical forces within the organism that require a sensitive observer as the measuring instrument to gauge them.

Clearly, this is a schism between two major outlooks in psychology. The tough-minded, empirically-oriented, pragmatic and (so it is alleged) characteristically American tradition in psychology has been behavioristic; it has been interested in elements rather than in wholes; and at first sight it fell in love with operationism. Its traditional enemy has been (to use William James's dichotomy again) tender-minded, not unwilling to flirt with intuition and the direct examination of phenomenal experience. In the second camp, the personality as a whole is an object of value, almost of veneration, and its warriors have had a deep distrust of attempts to fractionate the individual, even conceptually. This feeling has been matched by the suspicion in the opposing group of anything that smacked of the occult, mystical, or otherwise nonscientific.

The two points of view are not wholly antagonistic or incompatible; neither are they as internally homogeneous as Murray depicted them. Despite his leanings toward a literary, historical, and intuitive approach, Murray was able to avoid the methodological pitfall of committing himself to the so-called ideographic method. He saw clearly that intuition had to be used systematically, even quantitatively, that the riches of personality, subtle though they were, could be abstracted in terms of a set of analytical ("nomothetic") variables so that general principles and laws might be discovered.

In 1943, Murray left the Harvard Psychological Clinic to play a decisive part in the work of selecting personnel for the Office of Strategic Services. Here was a selection task that presented an unusual challenge: people had to be chosen to become spies, saboteurs, operatives of all kinds in a complex undertaking of intelligence, psychological warfare, and related matters—much of the work being done under the most hazardous of overseas conditions. At half a dozen centers, 5391 recruits were studied intensively by teams of psychologists, psychiatrists, and other social scientists under the leadership of Murray and a group most of whom had had some experience at the Harvard Psychological Clinic. In *Assessment of Men* (OSS

Assessment Staff, 1948), which reports on that research, a debt is acknowledged to the work of Simoneit and other German military psychologists and to the British psychiatrists and psychologists who devised and conducted the War Office Selection Board program for testing officer candidates in the British army.

Yet when all these acknowledgments are made, the work that is reported in *Assessment of Men* clearly bears a family resemblance to that in Murray's earlier collaborative volume.

The procedural steps outlined in that report had a good deal of influence on the design of our own research. They had many points in common with the "elementaristic" techniques of the AAF psychologists (not yet published when our work began). The main differences are: (1) The abstract variables rated by the OSS group were thought up from theoretical and common-sense considerations instead of being extracted statistically (by factor analysis). (2) They were measured on the basis of elaborate assessment techniques instead of objectively scorable tests. (3) The job analyses made by OSS were speculative rather than systematically empirical. Perhaps an even more significant difference is (4) the failure of the OSS system to specify validation against a criterion, followed by revision of the assessment methods and further cross-validation. (5) What the OSS procedure included, and the AAF omitted, was individualization: an attempt to trace and respect the unique individual pattern of the variables.

In point of fact, empirical refinement through constant checking against a criterion was simply impossible in the OSS set-up. There was no sufficiently meaningful criterion immediately available, such as success in a training course. Graduates of the assessment program went on clandestine assignments to remote and inaccessible parts of the world. A more difficult situation for getting a good criterion could hardly be conceived. Yet the assessors made persistent efforts to get whatever criterion data they could, and they present their modest validities with full consideration of the faults in the criteria. In general, the correlations run from .08 to .53, averaging about .25.

If this seems a sorry performance compared to the AAF validities, remember not only the vastly easier circumstances for research under which the latter were obtained but also that they are against a simple criterion—that of passing or failing in stateside training schools. When the Air Force psychologists tried to get measures of performance in the field—e.g., how well the fighter pilots made out in

combat—they found many of the same difficulties that plagued the OSS teams. Nevertheless, persistent effort turned up a variety of possible criteria: number of enemy aircraft destroyed, promotions in the field, decorations, flying accidents, overseas efficiency ratings, etc. *Not one of these criteria correlated with any measure derived from a test at a level significantly above zero in any combat validation study!* In a way, this is not surprising, since the selection measures had all been geared in so closely with the very different tasks and situations of training. Nevertheless, the OSS validities look rather better when one considers that they are based on field criteria of the kind that the highly touted AAF methods were totally unable to predict.

#### THE UNIVERSITY OF MICHIGAN PROJECT ON THE SELECTION OF CLINICAL PSYCHOLOGISTS

A direct and immediate descendant from the *Explorations* studies through those at OSS was the VA project on the selection of clinical psychologists, published under the title *Prediction of Performance in Clinical Psychology* (Kelly and Fiske, 1951), and known more familiarly as "the Michigan project." While the OSS work was being written up, a number of its principal figures took an active part in putting into effect one of the recommendations they were to publish in the last chapter of *Assessment of Men*: to apply the same methods and principles to the selection of professional personnel. Of the persons most immediately responsible for the report on the OSS work, all but Murray himself—Donald Fiske, James Miller, Donald MacKinnon, and Eugenia Hanfmann—were intimately involved in the new project, and there was a total of sixteen persons in all who served on the staffs of both enterprises.

The work on the selection of psychologists started in September 1946 and lasted for five years. Thus, in a sense our own work began earlier and lasted longer; but, effectively, the Michigan Project had a year's head start with a large and very active staff. For that reason, work that was in some sense coterminous with our own appeared to us in the guise of previous research, from the design of which we could learn while setting up our own. Throughout the four years that followed, there were fairly frequent exchanges of ideas and information, with several conferences and a fair amount of correspondence as well as the exchange of progress reports. The extent of our indebtedness to this work is difficult to estimate, but it is certain

that the Topeka study would have been a different one had the Michigan project not existed.

Kelly and Fiske point out one major difference between their project and the assessment program of the OSS: In the work on clinical psychologists there was "much less emphasis on qualitative formulations . . . as contrasted with the emphasis on ratings"—quantitative ratings made by individual assessors, as well as by groups, and covering descriptive as well as predictive variables. The major feature that these two projects shared was their heavy emphasis on *situational tests*: attempts to set up in an experimentally controlled way situations (usually stressful and demanding ones) in which candidates could be directly observed struggling to cope with problems assumed to be miniatures of those they would encounter in the work for which they were being selected or, at least, directly relevant to it. The emphasis changed from the attempt to create a vocational miniature to an effort to set up situations that would bring out diagnostically revealing aspects of personality.

The Michigan Project was set up in such a way as to make possible direct comparisons of free clinical ratings and the kind of objective test scores that were the principal predictors in actuarial studies, though it did not provide any direct comparisons between predictions from clinical and statistical analyses of the same data. Finally, perhaps the other most relevant aspect of the Michigan Project was its heavy emphasis on the development of criterion measures.

The Michigan research thus influenced ours in two principal ways: We considered that it was useless for us to duplicate the extensive work they were doing in trying out objective and situational tests, and we profited from their generosity in keeping us in touch with the variety of criteria that they tried out.

#### DAVID RAPAPORT AND DIAGNOSTIC PSYCHOLOGICAL TESTING

A somewhat separate stream of influence on our project came from the teachings and the research findings of David Rapaport. Rapaport had an important hand in the framing of the original experimental design, since he was the first principal investigator and carried the major work of the project himself for its first year. Less obvious is the host of ways in which his approach to clinical research pervades the entire work of the project. Both of the authors were

trained in diagnostic testing by Rapaport himself and by his students.

We wish to call attention here to certain of his principles and points of view toward clinical practice and research, rather than to findings that might be compared to other studies (see Appendix 7.1). Rapaport's thinking was affected a great deal by *Explorations in Personality*, yet when he read it he was already using psychological tests to diagnose psychiatric patients in a way that emphasized detailed understanding of the processes of response through an intensive inquiry into test rationale—an attempt to think through the nature and significance of tests in terms of theory of personality (the psychoanalytic theory of thought-organization).

This emphasis on test rationale leads to a type of testing that is never "routine." Giving a psychological test is always an experiment, always an opportunity to observe in action the complexities of thinking as they have been conceptualized by psychoanalytic theory. The attempt to understand in detail the processes that give rise to each feature of the test performance leads to a repeated examination, both qualitative and quantitative, of the verbatim text of the testing performance—the behavior and attitudes of the patient and his relationship to the tester, as well as his verbalizations.

Rapaport's work emphasized a relatively neglected aspect of test performances in order to learn the most about a side of personality that was also being neglected in the rising enthusiasm for psychodynamics. The careful study of *formal aspects* of responses, in addition to their content, made it possible to discern the operation of characteristic modes of adaptation and defense organized into ego structures. In terms of psychoanalytic theory, this meant using not only concepts of psychosexual development, symbolic representation of conflicts and complexes, and the defense mechanisms, but newer contributions to ego psychology, such as the adaptive point of view (Erikson, Hartmann; see Rapaport, 1956), the conflict-free sphere of the ego (Hartmann), along with Freud's concept of primary and secondary processes and his description of the latter in terms of concept formation, attention, concentration, anticipation, etc. From Rapaport we learned how characteristic types of thought-organization and impulse-management operated in the manner of structures, setting limits within which familiar dynamic principles operated.

Finally, the research that was published in the two-volume monograph, *Diagnostic Psychological Testing* (Rapaport et al., 1945-46;

see also Schafer, 1948) underlined the principle of the battery of tests. The idea of the battery is simply that a variety of tests pose a number of quite different types of tasks to the subject, differing, for example, in the structuredness of instructions and materials, thus requiring him to show the range of adaptive and defensive resources at his command. In terms of such a comprehensive inventory of adjustive techniques, analyzed in the light of the ego-structural point of view mentioned above, diagnosis was something far more meaningful than selecting a pigeonhole in which to categorize a patient. It became a detailed assay of personality structure, emphasizing types of adaptation and ways in which particular aspects of an individual style of adjustment had decompensated to produce pathological and symptomatic manifestations in various realms of function. This kind of personality assessment played a large role in the research to be reported.

#### CLINICAL VERSUS STATISTICAL PREDICTIONS?

The two traditions we have described, the clinical and the actuarial or statistical, might be thought of as supplementing or complementing each other, and thus as existing together in a co-operative relationship. Actually, however, advocates of the two approaches often took an intensely competitive attitude toward each other. The advocates of the clinical or case-history approach claimed that the sensitive clinician, using his trained intuition, could seek out data that would be important for each unique case and could discern special patterning of variables without which truly accurate prediction could not be made. Proponents of the statistical approach maintained that the objectivity and mathematical sophistication of their method enabled them to find and use the optimal weighting of whatever could be reliably measured to predict any definite criterion.

It is only fair to the authors of *The Prediction of Personal Adjustment* to note that they had enough perspective to maintain a co-operative, rather than a competitive, point of view. Despite their expert familiarity with the actuarial approach, Horst and his collaborators were sympathetic to the "case study" approach, as they called it. They pointed out ways in which careful analyses of individual cases play an indispensable role in predictive research and were willing to suspend judgment on the comparative merits of the two types of method until more research should come in.

Murray, on the other hand, phrased the issue as a conflict between organismic and elementaristic principles of assessment. From today's perspective, this reformulation did nothing to clarify the issues; we mention it to point out the fact that the problem looks different today than it did in 1946-47, when our project was begun. A good deal of experience has accumulated, there have been a number of empirical studies attempting to compare clinical and actuarial predictions, and much thought has been devoted to the logical issues.

The most lucid and important contribution to this controversy appeared after the first draft of the present book had been written. In *Clinical vs. Statistical Prediction* (Meehl, 1954), Paul Meehl examined the issues in such a thorough and stimulating fashion that his work cannot be ignored; in addition, he made a valuable collection of the available evidence.

Let us consider the theoretical argument first. With the work of T. R. Sarbin—both his research (1942) and his theoretical analysis of the problem (1944)—the controversy entered a new phase. Sarbin maintained (as had Lundberg, 1941) that there is no logical difference between the two types of prediction: under any flag, prediction is essentially actuarial in structure, inferences about future behavior always being probability statements based on past experience with similar classes of cases. The apparent difference, Sarbin maintained, in brief, came about only because the clinician refused to recognize that he was operating unconsciously in an actuarial fashion (consequently doing a second-rate job of it) and put up a smoke-screen of talk about uniqueness, pattern-analysis, and the like.

In essence, Meehl's rejoinder maintained that there are two essential differences between the two methods: *First*, the clinician does not make his prediction solely from ordering a case to a certain class of similar cases, defined by available data and having a predetermined average frequency of observed success. Instead, he uses the data to formulate a structural-dynamic hypothesis about the personality and gets his specific prediction by applying a psychological theory about the way a personality, so constituted, would react under the given circumstances. Logically, it might be possible to set up an actuarial predictive scheme that would incorporate transformation rules based on the principle of intervening variables, but it would fail to operate as a clinician does because of the *second* essential difference: An actuarial system leaves no room for judgment and, hence, for errors of judgment; but this means that ultimately the data it

works with would have to be defined in objective terms. The laws of behavior incorporated into it would have to be formulated in concrete terms, encompassing all the types of meaningfully equivalent stimuli that call forth a similarly multiform class of equivalent responses.

Let us consider a very simple example. Suppose that we accept, for the purposes of argument, the proposition that frustration produces aggression, and that we can therefore predict aggressive reactions whenever a person encounters frustration. But what constitutes a frustration?—a huge class of situations; and just as huge is the list of possible forms the resulting aggression might take. As Meehl points out, one of the characteristics that differentiate a trained clinician from a clerical worker, who can apply a statistical formula or actuarial table, is the capacity to judge the psychological equivalence of externally different stimuli or responses and to recognize that a unique situation has elements of similarity with known cases.

We agree with Meehl that there is a difference in the logic of clinical and statistical predictions and that the former *can* operate more flexibly, broadly, and creatively so that it is capable, in principle, of yielding the superior predictive result. But this by no means implies, Meehl continues, that the clinician necessarily will do the better job whenever the two are compared, since his freedom is also a freedom to make many kinds of errors.

As Meehl presents it, therefore, the problem is to find the proper sphere within which both the clinical and statistical methods of prediction should operate. In seeking a solution, Meehl used not only logical analysis but also comparative empirical studies, collecting and reporting on twenty studies more or less relevant to this issue. Cronbach (1955) has added several others.

Before looking at the evidence on how clinical and actuarial predictors have performed, let us make sure that a few distinctions are clear. The two types of studies tend to differ in a number of ways that do not necessarily go together. A clinician usually predicts different kinds of things than a statistician does (*e.g.*, how a patient will respond to a certain type of psychotherapy, as opposed to passing or failing in a school). He does so case by case, generally using qualitative language instead of ratings or probability estimates. He chooses different types of data—interviews, projective tests, interactions with the subject, life histories, and the like, instead of objective tests and measurements. The essential point in these and other distinctions is

that clinical judgment plays a variety of roles in predicting behavior. It may enter inappropriately at several stages of the predictive process, attempting to substitute for acquaintance with the facts, at the levels of: job analysis (trying to find out exactly what is to be predicted), choosing intervening variables that must be assessed, or choosing types of data-gathering techniques that will give information about these personality variables. More functionally, it may enter in the gathering and processing of data to give measures of intervening variables or in the final step of combining the processed data to yield definite predictions. (These steps are discussed more fully in Holt, 1958a, where the present argument is developed in a slightly different form.) Any given study may be clinical in one or more of these respects and actuarial in others; thus, many hybrid types occur, and it is confusing to try to classify them all according to a simple dichotomy—clinical or statistical.

Consider, for example, the pioneer attempt made by a group of clinicians at Massachusetts General Hospital to predict success in primary flight training, before American entry into World War II (Finesinger *et al.*, 1948). They studied 150 naval flight trainees (35 of whom washed out) with over 150 measures derived from psychiatric interviews, a supplementary questionnaire, the interaction chronograph (a mechanical means of measuring how long each of two or more persons speaks in an interchange of remarks during an interview), measures of respiration, muscle tension, and heart rate; tests of the autonomic nervous system; the electroencephalogram; five psychological tests, most of them projective; and handwriting. An attempt was made to develop ratings and scores so that the psychiatric interview and projective tests could be objectified, but the reliability of the scores was never satisfactory.

Although the total set of validity coefficients when these 150 scores were correlated with pass-fail was an essentially random distribution centering around zero, there were so many scores that a fair number of high correlations resulted. None of them was based on a clinical evaluation of any sizable body of data. The experimenters proceeded to combine these into batteries that had multiple correlations with criteria from .7 to .9.

The five most promising-looking tests, yielding 32 scores, were tried out on a new group of 506 trainees. The results were a classic example of cross-validation shrinkage. Validities fluctuated drastically from one to another of the four subsamples involved, and none of

the multiple correlations was as high as .2 on more than three of the subsamples. Clearly, the original "results" had been capitalizations of chance findings, which did not repeat themselves.

In a number of ways, the study is an instructive example of how *not* to do selection research. The authors were clearly not familiar with the existing literature on the subject. They made no attempt at job analysis, apparently unperturbed by lack of knowledge of what it was they were trying to predict: "Although no detailed job analysis of flying was available to the investigators, *it is clear from general knowledge* [italics ours] that besides certain mechanical and manipulative skills, the job of flying and fighting a plane to a large extent involves skill in interaction, interaction not merely between the flyer and other members of his squadron, or other members of the ground forces, but also in fighting an enemy" (p. 45). Even the plausibility of this argument for use of the interaction chronograph (measuring aspects of conversational interaction in an interview) fades when one considers that skill in fighting with a plane plays no part in the criterion they used—pass-fail in flight training.

Was this a clinical or an actuarial study? Clinical judgment was used at many places—in place of a job analysis, as a way of choosing predictive instruments, in the digestion of data to provide scores—but not, oddly enough, at the critical point: making the final predictions. That was done by means of the strictly actuarial, statistical technique of multiple regression. The question therefore seems naive; in some respects the study was clinical, in others statistical. There is no such single approach as *the* clinical method or *the* actuarial technique.

Meehl failed to make this point, however, and unwittingly perpetuated the conflict he was attempting to mediate by continuing to speak as if clinical and statistical prediction could be meaningfully pitted against each other. From his discussion of the issues, it is clear enough that he was primarily concerned with the final phase of the predictive process and that he intended to classify a study as clinical only if clinical judgment were used, instead of some simple, statistically derived rule in producing predictions. If the process involved no judgment, no knowledge of psychology—just the capacity to follow directions carefully or to do some arithmetic accurately—it would be classified as actuarial.

Most of the studies cited by Meehl made some approximation to the design first used by Sarbin: An actuarial system is set up to

predict some such criterion as dropping out of freshman year in college; the data used in the statistical prediction (*e.g.*, high-school grade-point average and an aptitude test score) are made available to clinicians, along with other data, such as an interview. The clinicians make either pass-fail or quantitative predictions, and these are validated on the same group of subjects along with the actuarial predictions. It would not be particularly instructive to review the details of particular investigations; some were better designed than others, or they deal with slightly different kinds of predictors and criteria. Despite such differences, in all of the studies that have been pulled together by Meehl and by Cronbach, the actuarial predictions do at least as well as the clinical ones, often slightly better.

The study by Melton (1952) is a representative example. He had fourteen counselors forecast the freshman grades of 543 entering students, using high-school rank, the ACE intelligence test, an interview of forty-five minutes to an hour, and scores on three other tests and inventories. The actuarial predictions were given by a multiple-regression equation using the first two of these, the weights having been derived from another sample. On the average, the counselors did worse and, as individuals, eleven of them were less accurate than the equation. When a counselor made his prediction after having seen the actuarial prediction, he tended to add error rather than insight.

Meehl's conclusion is that clinical prediction is an expensive and inefficient substitute for the actuarial method, and one that keeps clinicians from using their talents more constructively when aimed at such a criterion as educational or therapeutic outcome. He believes that something in the nature of clinical intuition exists but its proper sphere is psychotherapy or exploratory research.

When one looks at the evidence from the position that we are urging, however, this conclusion is hardly tenable. If two attempts to predict the same behavior differ significantly in the role played by clinical judgment, as against actual study of the facts in one or more of the four earlier parts of the predictive process, a comparison of the successes of the two attempts can tell us nothing definite about the effectiveness of clinical judgment at the final, crucial stage. For this reason, the comparisons were not pertinent to the point in the studies Meehl cited. Particularly at the vital third step, the predicting statisticians have had the advantage of having previously studied the way

their predictive data are related to the criterion, and the clinicians have not.

For purposes of exposition, let us suggest a slightly different typology.

*I. Pure actuarial:* only objective data are used to predict a clear-cut criterion by means of statistical processes. The role of judgment is held to a minimum, and maximal use is made of a sequence of steps exemplified in the most successful Air Force studies in selecting pilots (job analysis, item analysis, cross-validation, etc.).

*II. Naive clinical:* the data used are primarily qualitative, with no attempt at objectification; their processing is entirely a clinical and intuitive matter, and there is no prior study of the criterion or of the possible relation of the predictive data to it. Clinical judgment is at every step relied on, not only as a way of integrating data, but also as an alternative to acquaintance with the facts.

*III. Sophisticated clinical:* qualitative data from such sources as interviews, life histories, and projective techniques are used as well as objective facts and scores, but as much objectivity, organization, and scientific method as possible are introduced into the planning, the gathering of the data and their analysis. All the refinements of design that the actuarial tradition has furnished are employed, including job analysis, pilot studies, item analysis, and successive cross-validation. Quantification and statistics are used wherever helpful, but the clinician himself is retained as one of the prime instruments, with an effort to make him as reliable and valid a data-processor as possible; and he makes the final organization of the data to yield a set of predictions tailored to each individual case.

If we now re-examine the studies compared by Meehl and Cronbach, we see that most of them have pitted approximations to Type I actuarial predictive systems against essentially Type II naive clinical approaches. It seems hardly remarkable that Type I has generally given better results than Type II; indeed, the wonder should be that naive clinical predictions have done as well as they have, in a number of instances approaching the efficiency of actuarial systems.

Other studies cited have come closer to comparing Type II with Type III—naive *vs.* sophisticated clinical predictions instead of clinical *vs.* statistical. For example, the prognostic studies by Wittman (1941) compared predictions of reaction to shock treatment, made in a global way at a staff conference, with a system she devised. But her system used highly judgmental predictive variables, as Meehl

himself points out, and they were combined using a set of weights assigned on judgmental, not statistical grounds. What she showed was that a systematic and comprehensive evaluation of the thirty items in her scale (all based on previous empirical work) gave a better predictor of the outcome of shock treatment than global clinical judgments, not so organized and guided. A study of movement in family case work by Blenkner (1954) came to a very similar conclusion with somewhat different subject matter. When social workers rated an initial interview according to their general impressions, they were unable to predict the outcome of the case, whereas when their judgments were organized and guided by means of an outline calling for appraisals of five factors which had been shown in previous studies to be *meaningfully*, not statistically, related to the criterion, then these judgmentally derived predictive variables, combined (like Wittman's) in an *a priori* formula, predicted the criterion quite well. Yet both studies were wrongly tallied as proving actuarial prediction superior to clinical.

In a panel discussion of Meehl's book, McArthur (1956) calls for studies in which the clinician and the statistician do not attempt to predict the same criterion, but each tries to find the kind of thing he thinks he can predict best, as well as using data of his own choice. He maintains that there have been, as yet, no studies in which the clinician has been given a chance to show what he can do on his own terms. In terms of this discussion, our research on the selection of psychiatrists is not a comparison of actuarial and clinical predictions on the same subject but mostly an attempt to give clinical prediction a maximal chance to show what it can do.

To do a sophisticated clinical predictive study (Type III), one ought ideally to be fully conversant with the methods of the actuary, even though one does not intend to imitate him in all respects. In the present project, we made the initial mistake of not studying the work of the actuarial tradition nearly as closely as we have done subsequently. We reasoned—correctly, we still think—that psychiatry is a very different job from, say, driving a streetcar, and one that makes its most important demands on the inner personality of the subject, rather than on easily measurable external abilities, attitudes, etc. The fallacious conclusion we drew from this premise, however, was that the people who had selected streetcar motormen had nothing to teach us. Reading about their work later on, we see that we could have saved ourselves time and headaches by a careful review of a for-

biddingly dull-looking literature. All of the issues look much clearer today than they did in 1946 and 1947. We groped toward Type III, learning the hard way. It is at least comforting, belatedly reviewing the actuarial literature, to learn that what we had to discover for ourselves had plenty of precedents.

## THE PROBLEM OF THE CRITERION

IN RESEARCH ON THE SELECTION OF PERSONNEL, THE "CRITERION" IS that *standard of performance* against which the soundness of decisions involving acceptance or rejection is later checked. Every such decision is implicitly or explicitly a prediction of future performance in some kind of work; therefore the criterion may also be looked on as *that which is being predicted*.

One tends to think first about the kinds of predictors to use and only later about the criteria against which their usefulness is to be measured. The experience of many investigators, however, and logical considerations, as well, point to the necessity of deciding on criterion measures early in the research—ideally, before predictors have been chosen at all. Obviously, there is no way of checking a prediction that it will rain two inches next week without some way of measuring a week's rainfall in a standard fashion.

This example from weather forecasting is cited to contrast the relative difficulties of measurement in a field such as psychiatry, where no universally accepted standard of excellence exists. Not only do we have the usual difficulties of getting precise and appropriate measures, but what we wish to measure is a subtle human performance of a kind that is hard to evaluate.

Since the final step in a selection study is the evaluation of re-

sults, the choice of a criterion determines the ultimate direction in which the research will move. Let us approach the choice of criteria by examining first the objectives of the Menninger School of Psychiatry. The official objective of the VA residency programs was to train psychiatrists to staff VA hospitals and mental hygiene clinics. But the aim of the program was never conceived so narrowly. Those in charge of training in the VA foresaw that residency programs would be worth a large investment of the agency's funds and efforts even if most trainees did not stay in VA employment. It was never made a condition of receiving a VA residency that a doctor should obligate himself for any specified period of service after training. It was plain that society would be served by training more psychiatrists, and veterans would receive a great deal of good care from these men during their training—not to mention the benefits from the fact that training programs attracted more and better staff men to VA hospitals.

The School's aim was, minimally, to select men who would (a) competently complete a three-year program (b) fulfill the requirements of the American Board of Psychiatry and Neurology for certification, and (c) remain in the practice of a branch of psychiatry. All three of these objectives imply pass-fail types of criteria, and criterion data of these kinds were not difficult to gather. We set up procedures to get these three items of information about each of our subjects.

From the beginning the faculty strove to train psychiatric leaders. This meant taking "only exceptional men"; and, judging by such statements in the early era of the Menninger School of Psychiatry, the only adequate criterion of the School's success, either in training or in selection, would have demanded a long-run follow-up, establishing which of the alumni became *leaders* in psychiatry—thinkers, teachers, innovators, writers, researchers, institutional chiefs, and powers in professional societies. Perhaps such a follow-up will one day be possible.

#### TYPES OF CRITERIA

The criteria we have mentioned so far may be itemized as *leadership*, *membership in a profession*, *technical or professional competence*, and *completion of professional training*. These need not be the only kind of criteria considered. One might also want to know how much personal gratification a psychiatrist—or psychiatric resi-

dent—will get from the work (an *adjustment* criterion); or what rewards of money, position, prestige, etc. he will receive as a psychiatrist (a *success* criterion); or what contributions to psychiatry he will make, by way of research, theory, or new diagnostic or therapeutic techniques (a *productivity* criterion); or what kind of work he will do in psychiatry (a *functional* or subspecialty criterion). Actually, we gathered data bearing on all these types of criteria, but the kind of criterion that seemed most important to aim for was a measure of competence. Neither success nor adjustment has the social importance of competence; productivity is one facet or measure of competence; and the functional criterion seemed interesting but not so much needed as a measure of competence (see Appendix 8.1).

#### IN-TRAINING VS. POST-TRAINING CRITERIA

The men were being assessed during the year before they began training (more than a year before in some cases), and it would have been possible theoretically to begin collecting criterion data on the day they entered training or to defer it until they had completed their careers. The later the date on which criterion data are gathered the better, no matter what kind of criterion is to be predicted. A psychiatrist may be obscure and impoverished, as Freud was during his first years of practice, yet later attain *leadership* and be considered a *success* by his colleagues. He may be slow to develop *competence*, yet reach a very high level of it after years of experience; he may make *contributions* to psychiatric theory or technique, the importance of which may not be recognized for some years. Just so, it may take him decades to find a niche and a way of life for himself in psychiatry in which he can find contentment and *vocational adjustment*. Finally, a psychiatrist has many opportunities during a long career to change the setting in which he works and his *type of practice*. Many have begun in institutional work, diagnosing and giving somatic treatment to psychiatric patients, then switched to private practice of psychotherapy with out-patients, ending up in a primarily psychoanalytic practice. In such a case, which would be the proper answer to the original question: What type of psychiatry will he practice?

The difficulty of predicting almost any kind of behavior increases with the lapse of time between assessment and criterion evaluations. This is an inevitable result of the way behavior is determined. It is always a complex function of determinants within

the *person* and influences, opportunities, and restraining conditions of his *environment*. We are on tenuous ground when we try to predict, from incomplete knowledge of only the person himself at one point in time, how he will act in a different (and thus unknown) situation at a future point in time when he has been subjected to an intensive experience, the very object of which is to *change* him! Moreover, many residents seek to change themselves via psychotherapy or psychoanalysis during or shortly after psychiatric training.

Furthermore, the settings in which graduates of the Menninger School practice different kinds of psychiatry are many and varied: One alumnus flies the rounds of Army clinics over the arctic wastes of Alaska; another hardly ever leaves his office-apartment overlooking Central Park in New York City; another has charge of a guidance clinic in a Southern city. These different locales influence what and how well people do in psychiatry, presenting as they do all degrees of opportunity, frustration, stimulation, isolation or interaction with colleagues, comfort or difficulty in nonprofessional aspects of living, and the like.

To complicate matters further, professional people often move about the country—and over the world—and are hard to keep track of. The graduates of the program scattered so far that it would have been impossible for a team of evaluators to visit them during a reasonable period of time. Even if this could have been done, we would have had to develop observational and evaluative techniques to enable us to assess in a comparable way the variety of functions performed by an alumnus who psychoanalyzes children in an office; another who works as part of a research team in a university hospital studying psychosomatic aspects of gynecological disturbances, another who is clinical director of a large hospital for psychotic veterans, and others equally diverse. Evaluations by colleagues would have been difficult to collect in adequate numbers (impossible in the case of men who work in professional isolation) and would hardly be comparable even if they could be obtained.

Therefore, the post-training follow-up was considered impracticable as a principal criterion, and we focused on criteria to be gathered during training. They were to be measures of competence, taken during the last two years of residency and supplemented by a modest amount of post-training information, including the *kind* of psychiatric practice established, obtained from the alumnus by means of questionnaires.

An in-training criterion may still be a valid measure of competence. The work of the resident, particularly in his last year, is mostly the care of patients; he is thus engaged in professional practice of a responsible kind. Furthermore, to be able to predict an applicant's ability to profit from psychiatric training would be an important and useful achievement. If at the time of application we could identify physicians who would be unable to complete residency successfully, we could prevent a great deal of personal tragedy and waste of time, effort, and money by physicians and by the training institution.

Although we did not attempt to predict choice of subspecialty, we did try to make separate predictions about different aspects of a given resident's work. This was done for two reasons: (1) if predictions of over-all competence did not prove to have validity, special skills might still be shown to be predictable; (2) experienced psychiatrists thought this kind of prediction valuable. In the end we obtained several sorts of criterion measures from the final years of residency, gauging over-all competence, competence in psychotherapy, diagnosis, administration, and therapeutic management, as well as several dozen measures of more specific functions.

#### DICHOTOMOUS VS. SCALED CRITERIA

The next decision to be made was whether to set up dichotomous (two-valued) categories of the pass-fail type or to attempt to measure a gamut for each skill.

There are some strong arguments for the simpler solution. It is easier and less time-consuming. The task of selection itself can be seen to be one of putting applicants into two categories: those who are acceptable and those who are not. And no one would have to strain to force what may seem to be essentially qualitative distinctions into a quantitative mold.

The basic objection is that *the cutting-point of minimal acceptability must vary from one time and place to another*. There are four main factors that determine such a cutting point: the social need for psychiatric services, the availability of training facilities, the number of men seeking training, and the stage of development reached by the profession in the transition through which it is going. Each of these four factors can affect the number of distinctions that can most usefully be made on a scale of work performance; they also

help to determine where on this scale a point of bare acceptability should be set.

During the time of this study, the demand for psychiatric services was exceedingly great. It is still growing. Many more psychiatrists will be needed in the next few years than there are physicians seeking psychiatric training. There is, therefore, little social need for means of drawing fine distinctions between degrees of excellence in psychiatry at the top of the scale, but there is a great need for means of making the basic distinction between *adequacy* and *inadequacy*. If there were many more men clamoring to become psychiatrists than there were effective demands for their services, it would be more urgent to be able to distinguish the highly competent from the merely adequate. Looked at from this angle, the problem of recruitment appears greater than the problem of selection, and the cutting-point in selection should be set as low as possible to maximize the number of practitioners, excluding only the most dangerously unfit.

Nevertheless, the problem of selection for psychiatry will never disappear, no matter how great the demand and the availability of training facilities in relation to the number of applicants. Psychiatrists, like other medical specialists, are trained by means of practical work with patients. It is, therefore, not possible to accept all applicants for residency training. For the protection of the young doctor and his patients, it is vital to bar from psychiatric training not only the incompetent but also those who seem likely to become seriously upset by psychiatric work.

Turning to the second factor, the limited facilities for psychiatric training, there are likely to be more applicants than there are openings in sought-after training centers, although residencies may go unfilled elsewhere. This was the situation at the Menninger School of Psychiatry throughout the time of the project. The problem of selection looks somewhat different from this vantage-point and becomes more nearly one of using the limited training facilities with the greatest efficiency, which means trying to accept only those who can complete training with profit, even at the cost of not admitting some good men. In such a situation, a cutting-point is needed at a somewhat higher level than that of minimal adequacy.

The numbers of applicants for psychiatric residencies increased greatly and rather suddenly after the war, creating new problems of

deciding where the line should be drawn between the acceptable and the unacceptable; any future change in the numbers of applicants relative to the social need and training positions will bring about a similar raising or lowering of the cutting-point, both generally and for any particular institution. The hostilities in Korea and the accompanying partial national mobilization, which could hardly have been predicted, greatly reduced the number of applicants for specialty training.

The psychiatric profession is changing profoundly. In just a few years psychiatry has moved out of the "lunatic asylum" into the modern hospital, the outpatient clinic, and office practice. We are beginning to see the employment of psychiatrists in industrial plants, by labor unions as well as by management, in school systems, in public health departments. Psychiatric epidemiology is in its infancy, but we hear from some quarters that the next great advances in psychiatry will be in prevention. The public is starting to awaken to the need for research in the field of mental health, and the great foundations have begun supporting such work in a way that may make careers in psychiatric research attractive to many. Such changes may affect the number and caliber of men attracted to specialization in psychiatry.

Shifts in the availability of training facilities, in the numbers of physicians who want psychiatric training, in the public's ability to pay for psychiatric care—these and many more changes may alter substantially the pressure on psychiatric training centers to accept a greater or smaller proportion of applicants. Persons who have the practical job of selecting residents next year or in another state may therefore need quite a different cutting-point from that in use here and now. The more finely divided the scale of valid predictions that may be made, the greater becomes the range of possible future cutting-points in practical application of the predictive methods. It seemed to us wise, therefore, to use whatever simple success-failure criteria were available locally at the moment but also to supplement them by the attempt to predict on a more finely differentiated scale and to obtain criterion measures on a comparable scale. (For further discussion of these issues see Appendix 8.2.)

#### RATINGS VS. SPECIAL MEASURES

How may competence during residency training be measured? Several possibilities presented themselves: ratings by supervisors, rat-

ings of the subjects by each other, formal written examinations on the content of the curriculum, special observations, and work samples. All of these except the last were actually used. (For the reasons we abandoned work samples see Appendix 8.3.)

These techniques all involve rating at some point or other, for there seems to be no way to get a meaningful measure of any important aspect of psychiatric work without using a human being as an instrument. Purely quantitative measures of work output might have been collected, such as the number of diagnostic summaries written up by a resident during his stay on a service or the number of patients discharged as improved who had been under his care. Study of such measures under the conditions of the residents' work quickly shows that they have little significance. By doing a sloppy and superficial job, one man may keep his paper-work completely up to date, while another, more thorough and conscientious, is behind in completion of reports but does a better total job. Similarly, residents may differ in the numbers of patients discharged who have been under their care because they have different standards of "maximum hospital benefit." It seems at first thought that reliable and valid measures of improvement applied by someone other than the resident would form ideal means of estimating psychotherapeutic competence. But no such measures existed and to develop them would have been a major research project in itself. Furthermore, adequate measures would have had to take into consideration the kind of patients under any given resident's care; for some residents had heavier and more difficult therapeutic assignments than others, and sometimes the best men would be deliberately given the most refractory patients. It is hard to see how adjustments in the criterion could be made without the intervention of judgment, which blurs the line between such a measure and a rating. For practical reasons, however, it was impossible to give this kind of detailed consideration to each resident's total case load, though our criterion ratings of psychotherapy involved a good deal of this kind of judgment, carried out informally and unsystematically.

Even *examinations* must be graded, which is a rating process, except for multiple-choice and other objective types of questions. A written examination (with essay questions) was tried by the School during its first year, as a way of objectively determining the readiness of some of the men to pass to advanced status. The first student body was quite heterogeneous, including some men with years of military

psychiatric experience, others with some residency credit from non-military sources, and some who had come directly from medical school; yet at first all had to take the same courses. Later, when the faculty was able to offer advanced as well as elementary instruction, it was by no means obvious who was ready for it. So it was announced that residents who wished to be considered for advanced standing would be examined.

We have the grades for twenty-eight residents who took the first examination and grades for the entire resident group on another such examination given in 1953. (The second examination is discussed in Appendix 5.2. See also Appendix 8.6.) The attitudes of many residents and even a few members of the faculty were too much opposed to academic examinations for us to have been able to use them as a major source of criteria. More important, none of the grades from examinations showed any consistent relationship to other criterion measures. It seems unlikely, therefore, that they measure the kind of competence in which we were interested.

We put our main reliance on ratings based on general observation of routine work, made by both the residents' supervisors and their peers. These formed our principal criteria, yielding scales that could be used for a number of different statistical comparisons. In most respects they turned out to be quite satisfactory, though they too required a good deal of developmental labor and went through several metamorphoses.

#### THE CRITERIA ACTUALLY USED

Out of all the possibilities we have been discussing, the criteria adopted were as follows: (1) completion of three years of psychiatric residency training; (2) entry into actual psychiatric practice; (3) certification by the American Board of Psychiatry and Neurology; (4) apparent "success" in psychiatric practice; (5) estimates, rated by fellow-residents, of level of competence during residency, in over-all performance and also in each of several major aspects of the resident's work; and (6) the same, as rated by supervisors.

1. *Completing three years of psychiatric residency.*—“Minimal professional competence is usually assumed on the basis of such a priori criteria as completion of professional training, passing a certifying or licensing examination, or continued successful practice” (Kelly and Fiske, 1951). There are a number of objections to the simplest

of these criteria—graduation from the program; considerations that make it quite different from the comparable criterion (in medical selection) of achieving the M.D. or (for clinical psychologists) of getting the Ph.D. The majority of psychiatrists probably start training right after completing an internship and take the three years consecutively. A very large number do not, however; there is no "statute of limitations." A man may drop out during his first year, seemingly giving up the intention of practicing psychiatry, and after several years resume training somewhere else, as one of our subjects did. Or, like another subject, he may hold such an important institutional position that he can be spared for only a year at a time and has to spread his training over a period of many years. In the last example, the man concerned was rated very highly by supervisors; yet if the present criterion were used quite literally, he would be a "failure" since he had not completed his three years by the time the project was over. An even greater degree of indefiniteness was lent to this criterion by the fact that the American Board of Psychiatry and Neurology decided to allow partial residency credit for psychiatric work in the armed services. The Board did not release any definite rules concerning how many years or months of training a man with military credit needed, however, so that in many cases the only final adjudication was the decision of the Board to allow a man to take its examination.

A number of men without previous residency credit did not complete three years in the Menninger School of Psychiatry, sometimes leaving despite the desire of the faculty that they stay. Personal financial considerations, the desire for greater variety of training experience, and, perhaps most important, the more immediate availability of psychoanalytic training elsewhere were some of the reasons that good men left the program early.

Usually there was evidence in the School's files enabling us to tell whether a man who had left had done so despite the School's wish that he stay, or because he was ordered (or more informally asked) to leave because of incompetence, unethical behavior, etc., or because he had to drop out in order to seek therapy for his own emotional problems. Nevertheless, such a distinction could not always be made with complete certainty.

Our mail follow-up of alumni shows that the vast majority of those who left finished psychiatric training elsewhere. Only about 7 per cent of those entering the Menninger School of Psychiatry

failed to complete residency somewhere. "Completing three years of residency" is therefore difficult to use as a pass-fail criterion since almost all subjects passed it.

The great majority of subjects who did not complete training failed to do so because they did not begin it; 26 per cent of rejectees made no effective effort to enter psychiatry after being refused by the Menninger School of Psychiatry. These men, so discouraged by rejection at the Menninger School of Psychiatry that they did not apply elsewhere or were rejected in other attempts they did make, form a rather important criterion group. Unfortunately, there is a degree of contamination from the selection process itself that is unavoidable. In some cases, the Admissions Committees felt that their responsibility to the profession or to the applicant himself dictated that they try actively to discourage him from entering psychiatry. Sometimes the letter of rejection referred to personality problems which would have to be resolved by treatment before the applicant should consider trying to get a residency again; in other cases, the reference to therapy was omitted and the recommendation was more categorical. It is not surprising that these letters were effective in discouraging aspirants, although in a few instances the recipient of such a letter persisted and did get training elsewhere. In a number of other cases, applicants who were told merely that the limited number of places had been filled by slightly better qualified men nevertheless gave up the attempt to go into psychiatry. It seems to us safe to assume that they were so poorly motivated for a psychiatric career, or so much aware of their own lack of fitness for it, that the decision not to go on can be used as evidence of "failure." We decided, therefore, to use these non-completers of training (or "Drop-outs") as a principal criterion group for all purposes except validating decisions of the Admissions Committee. The degree of contamination is much less when Drop-outs are used to validate interviewers' or testers' ratings, and nonexistent when the predictor is a test score of some kind or a rating that had no influence on acceptance, such as one made by one of the experimenters.

For these reasons, the criterion of *failure* to complete residency anywhere may be used as a valid criterion, but it needs to be supplemented.

2. *Successful establishment of a psychiatric practice.*—This, too, is a criterion of limited usefulness. Every subject who had some training in the Menninger School of Psychiatry, completing it there or

elsewhere, succeeded in establishing a psychiatric practice,\* and most of those who had not yet completed training at the time of our last information about them were filling useful posts in psychiatry. Among rejectees, many more did not meet this test (24 per cent were not practicing psychiatry), so that it does seem to have some usefulness, again at the extreme low end of the distribution. Most persons in this category, incidentally, discovered either before or during psychiatric training that their interests and abilities lay in other specialties or in general practice, to which they turned usually with apparent success.

In practice, there was little difference between our first two criteria—completing training and establishing a psychiatric practice. Only a few subjects who failed to finish training attempted to practice psychiatry anyway (though this is legally possible), and fewer still kept it up for more than a year or so. The noncompleters of residencies constituted too small a group to be very useful, and since they were almost all included in the larger body of "Drop-outs" from psychiatry, the second criterion was almost always used in preference to the first.

3. *Certification by specialty board.*—In many respects, attaining diplomate status is a criterion of successful completion of psychiatric training. Psychiatric residency programs are set up in their existing form largely in order to satisfy requirements of the American Board of Psychiatry and Neurology. There are three kinds of examinations given by the Board corresponding to three types of certification: in neurology, in psychiatry, and in a combination of these. (For a further description of this criterion see Appendix 8.4.) The principal immediate goal of most residents is to obtain the diploma of the Board. There are some qualifications to be considered, however, in appraising the usefulness of certification as a criterion for our purposes:

a) Some members of the profession express dissatisfaction with the standards established by the Board. They disagree with the emphasis given to various aspects of psychiatry (about which representatives of different psychiatric viewpoints will inevitably disagree), and the general level of competence required is not high enough to command universal respect. Moreover, one may justly question whether it is possible to ascertain a man's fitness to practice as a specialist on

\* This statement applies only to our experimental group, not to all residents. See "General Nature of the Research Group," Chap. 4.

the basis of a six-hour oral examination, no matter how well done. There would be no widespread disagreement, however, with the proposition that a competent psychiatrist can or should be able to attain diplomate status if he really wants to. Stories are often told of outstanding figures in the field who failed or were "conditioned," but they all end with the psychiatrist's ultimately getting the diploma.

b) The fact remains that some psychiatrists do not seek (or in a given time interval are not successful in obtaining) certification by the Board, and that among them are many persons who fill quite adequately their professional roles. It is difficult to estimate the amount of error that this group introduces; in the light of the social (and in some cases, financial) pressures to obtain the Board's diploma, however, it cannot be very large, especially if failure to pass on one's first or second attempt (with eventual success) is not counted heavily against a man.

c) A more serious difficulty is the fact that two years' experience is required in addition to the residency before a physician may take the Board's examination. Add to this some additional delay for normal procrastination and the unforeseen contingencies that may cause a person to delay an optional ordeal, and it becomes apparent that not being certified during the first three or four years after residency cannot be held against a man. This means that the criterion of certification could not be used during the period of our research for the last two classes, which played a crucial role in the final research design. What might have been a main criterion measure was therefore reduced to a secondary role.

Since the list of new diplomates is published in the *American Journal of Psychiatry* a few months after each semiannual examination, collecting the criterion data is a simple matter. It is less easy to ascertain the names of those who fail; we finally decided to get this information from the follow-up questionnaire but found it of little value.

4. *Apparent "success" after training.*—Another criterion measure was derived from the follow-up questionnaires, thus dating from the first to the third year after completion of training. It is an estimate of professional "success," judged from the information yielded by the questionnaires in their entirety. (For details on how this judgment was made see Appendix 8.5.)

5. *Evaluations of the residents by each other.*—Peers' Evalua-

tions, as we called them for short, were part of Dr. Rapaport's original plans for the project. Although they offered their own special difficulties, we relied on them as a principal criterion second only to Supervisors' Evaluations.

There were reasons in the nature of the School's working situation to expect residents' evaluations of each other to have a good deal of validity. On most services, residents worked together closely, sometimes consulting each other on problems they might not bring to supervisors. At times they worked in close contact for longer periods than the duration of any one supervisor's contact. And occasionally the residents had more intimate knowledge of each other's work than did a particular supervisor.

For example, a resident once told us that when he came onto a service and took over patients who had been under the apparently satisfactory care of one of his classmates, he found that there were deficiencies in some of the diagnostic work-ups and difficulties in the previous resident's relationships with some of the patients. The first man had been able to conceal these faults from his supervisors, but his successor had to pick up the pieces he left and could hardly help learning a good deal about his work.

On the other hand, although most residents were personally acquainted with a large percentage of the total group, they were not likely to know the work performance of as many residents as each supervisor knew. Their standards for evaluating each other were perhaps less mature. Because there were more close friendships within the resident group than between residents and supervisors, it might be expected that personal likes and dislikes would interfere more with the former's evaluations than the latter's. It is clear, however, that these two types of evaluations have a lot in common; and, as it turned out, Peers' Evaluations correlated fairly well with Supervisors' (see Appendix 8.6).

We collected Peers' Evaluations on three occasions. The first two were group meetings, convened for this special purpose; mimeographed forms were handed out and were completed on the spot. These forms began with a list of the Fellows currently in residence (each time including a few who had recently left), followed by a series of blanks to be filled with the names of the men whom the respondent considered the three "best all-around psychiatrists," "best psychotherapists," "best research psychiatrists," and "best administrators." In addition, the form for Peers' Evaluations I called for choices of

three on eight other aspects of work and personality, for indications of the respondent's "best friends," and for some miscellaneous information. (A copy of the forms for Peers' Evaluations I, II, and III, complete details about the way the data were gathered and some of their interrelationships are given in Appendix 8.7.) Peers' Evaluations I had been obtained in December 1947; six months later, when we were starting a second round of Supervisors' Evaluations, Peers' Evaluations II were collected in a similar way. The main difference was that instead of the additional types of choices, there was a space by each name on the list for a *rating* of over-all competence on a four-point scale (finer discriminations could be made also by the use of decimals). Since these ratings correlated fairly highly with those of supervisors (.66 \*), it was therefore less necessary to have both sets, so we concentrated on improved methods of obtaining Supervisors' Evaluations for the next couple of years.

When it came time to collect criterion data on the main group of subjects in the final predictive study, we decided again to get Peers' Evaluations. Again, a list of eighty-three residents likely to be known to those currently available was included, and the respondent was asked to rate only those whose work he knew personally. No sociometric choices were called for; instead, ratings on a ten-point (decile) scale on Over-all Competence, Competence in Psychotherapy, Diagnosis, Management, Administration, and Liking (that is, a respondent indicated on the same ten-point scale how well he liked his fellows as persons). The residents did a conscientious job. The resultant ratings were highly reliable.

6. *Evaluations by supervisors.*—In deciding to rely primarily on the judgment of the trainers to evaluate the proficiency of the trainees, we reasoned along the following lines: It is not the function of a research team to decide what constitutes professional worth; that is a matter for the judgment of mature members of the profession concerned. We found that, in practice, psychiatrists are evaluated by their colleagues, with the opinions of a man's senior co-workers or supervisors and those who have opportunities for the

\* Bold-faced type indicates significance at the 1 per cent point or better; correlations significantly different from zero at the 5 per cent point appear in italics. This code is used throughout the book. Unless otherwise specified, we have computed the significance of correlations using only one tail of the t-distribution, indicated by reference to the .01 significance *point*; reference to significance *level* indicates that a two-tailed test has been used (see pp. 361ff.).

closest observation of his work being accorded the most respect and weight.

Our main criterion procedure was based on these assumptions and facts. We attempted to work with them in a way that would be systematic and minimize as many sources of error as possible.

We believe, therefore, that the criterion with the greatest inherent validity is the pooled evaluations of qualified supervisors who have been in close contact with the resident's work in the later stages of his training.\* This type of judgment is usually the basis for hiring and promotion and for other indications of professional success.

Residents were, of course, evaluated by their supervisors directly for the purposes of the Menninger School of Psychiatry, but we found on careful examination of these qualitative and quantitative written reports that they were irregularly submitted and too scanty in their descriptions of the resident's actual work. By holding loosely structured interviews with each supervisor as often as we needed evaluations, we found that we could get the kind of data we wanted in fuller detail and, moreover, that supervisors were willing to be franker with us, knowing that what they reported was kept confidential and would not affect a resident's subsequent career.†

Consequently, we collected such evaluations in eight series, changing and developing our methods as we went along. The full details on how each set of data was gathered are given in Appendix 8.8, with explanations of the changes made. Here it is enough to note that we learned to provide more detailed instructions to help the supervisors furnish the facts we needed; we improved the rating scales, increased the number of variables rated and defined them more explicitly, and got better information on the extent, dates, and nature of the supervisory contact.

The first three sets of evaluations (*Supervisors' Evaluations I, II, and III*) were gathered at six-month intervals from the winter of

\* This is not to deny the possibility that these ratings contained a good deal of hearsay reputation nor that to a certain degree they may represent the amount of match perceived by a supervisor between the resident and a stereotype of what a good psychiatrist must look like. We have no way of measuring the amount of error in the criteria introduced by such factors, though we believe it to be small.

† See Chap. 16, which synthesizes the picture of a good psychiatric resident that emerged from an analysis of these qualitative evaluations. In a way, it is a fuller specification of what the *Supervisors' Evaluations* mean in terms of the kind of person and behavior that they denote.

1947-48 to that of 1948-49. Over-all Competence was rated on a four-point scale, which the supervisors were allowed to use more precisely by adding, at first, pluses and minuses, later decimal divisions. In Evaluations I only, we also got each supervisor to choose the three men he thought potentially best in administration, psychoanalysis, and neurology, respectively, and the five best and five worst in Overall Competence. Beginning with Evaluations III, we got *ratings* of competence in Psychotherapy.

The rest of the evaluations differed mainly in two respects: We collected ratings on many more aspects of work (all of the work functions about which we made predictions in the Predictive Study; see Chap. 10); and ratings were now made on *graphic* rating scales. Thus, there would be a line, marked with percentile indications, on which a mark would be made by a supervisor near the midpoint if he considered that the resident in question was average (at the fiftieth percentile of all residents he had known) on the particular work-function being rated, e.g., Diagnosis. Starting with Evaluations V, a supervisor would be presented with a sheet containing such a line for each resident in the class, on which his competence at Diagnosis was to be rated; and one such sheet for each major variable—Management, Psychotherapy, etc. The ratings he made at different times were recorded on the same sheet, so that he could make man-to-man comparisons and put all residents he supervised in their proper rank order.

We also made some changes over the years in the ways the ratings were combined to give a final consensus evaluation of each resident—changes that resulted mainly from the discovery that we got more reliable results by using only supervisors who had had recent and intimate contact with the resident's actual psychiatric work, rather than relying on large numbers of faculty members who might not know the man equally well. (Techniques of putting the ratings together are given in Appendix 8.8.)

#### THE NATURE OF THE SUPERVISORY GROUP

Professional persons making the evaluations were the members of the faculty who gave clinical instruction: psychiatrists and psychoanalysts, the latter group including a few psychologists of senior rank and long experience who took part in supervising the residents' psychotherapeutic work (see Appendix 8.9).

Of the 58 people who gave us evaluations, almost all were dip-

lomates in psychiatry or neurology, and a few were diplomates in both. Only 8 of the 58 had had less than three years' experience post-residency. There were 23 full-time VA employees and about the same number of consultants or attending physicians at Winter VA Hospital, on the wards where the residents worked. (The rest supervised those parts of residents' work that was done at the Menninger Clinic or Southard School.) Almost all of the group served during more than one period of evaluation. Of these, 30 gave us 3 or more sets of evaluations, 10 gave us at least 5 sets, so that most of the group was quite used to the task.

Over 80 per cent of the supervisors who served as raters were more or less closely identified with psychoanalysis: 37 were candidates in the Topeka Institute for Psychoanalysis (during the period they were giving us evaluations, most of these were practicing as general psychiatrists, not analysts); 6 were graduate analysts; and 6 more were training analysts. The rest of the evaluators were general psychiatrists, administrators, or neurologists—4 were in the last group.

It is a truism of good scientific method in any predictive study to keep the predictions separate from the criteria. By the circumstance that 10 of our supervisory group did some initial interviewing, we ran the risk of such contamination. The chances were against a particular interviewee's ending up working on the service of his interviewer, and in the few cases where it happened, we excluded this supervisor's criterion evaluation of the man in question. Therefore, none of the validity figures is vitiated by this particular type of contamination (see Appendix 8.10).

#### RELIABILITIES OF THE CRITERIA

In order for judgments of professional competence to be valuable, they must be reliable: The judges should agree with each other in rating competence (tested by coefficients of internal consistency) and give the same ratings if the procedure is repeated later (tested by coefficients of stability). (Full details on the reliabilities of all criteria are given in Appendix 8.11.) Here it may suffice to say that all of the ratings were internally consistent enough to be usable, and for the most part they stand up under statistical scrutiny very well. The least reliable were Supervisors' Evaluations II (Over-all Competence) and III (Psychotherapeutic Competence); the reliability of the mean rating of each was .66, which is only fair. Of the seven other principal measures of these two kinds of competence, however, three

were between .8 and .9 (good reliability), and four over .9 (excellent). It is a rare study in this field that can produce higher coefficients of internal consistency for its criteria. Coefficients of stability are somewhat lower (from .63 to .77 for lapses of six months or a year) but satisfactory, considering that performance itself probably changed at least as much as opinions.

#### FACTOR ANALYSIS OF THE CRITERIA

The residents and the supervisors were called on to make highly differentiated ratings, covering many aspects of psychiatric work. Implied in our getting so many measures was the assumption that they were not perfectly correlated, and, therefore, it might be possible to predict different types of competence.

We intercorrelated the Peers' Evaluations, and also the Supervisors' Evaluations to see how much distinction was, in fact, being made. The criterion variables all correlated with one another so highly (see Tables 8.15 and 8.16 in Appendix 8.12) that we suspected there might not be much independence between the ratings. Consequently, we subjected the two sets of correlation to factor analyses. (For details of these analyses see Appendix 8.12.) The results were quite simple: There was only one factor in the Peers' Evaluations and only one of any significance in the Supervisors' Evaluations. This means that Over-all Competence is the only criterion in which one can have a great deal of confidence; the ratings of other variables are for the most part minor distinctions within the global judgment of general competence. There are two possible reasons for such a finding: Either the judges' observations did not permit them to distinguish one kind of competence from another or, in reality, only one major type of competence was required of residents in the Menninger School of Psychiatry (and a good diagnostician was, for example, almost always a good psychotherapist and ward administrator). Indeed, there is no way to tell just how highly correlated a set of perfect measures of these skills would be. It seems likely to us that the work-functions of the psychiatrist require skills that are actually highly correlated.

#### THE QUALITATIVE CRITERION

The basic task of selection is to separate acceptable from unacceptable prospective residents. Consequently, we decided to have

a dichotomous or (as we called it) Qualitative Criterion, based on Supervisors' Evaluations of competence in training.

Using both the quantitative ratings and the qualitative discussions of each man's work, together with all additional information from other reliable sources, we first divided all subjects who had entered the Menninger School of Psychiatry into two groups: Satisfactory and Unsatisfactory. We then further split them into five subgroups, as follows: The Satisfactory residents were divided into the "Superior" (those receiving the highest 20 per cent of the ratings) and the merely "Adequate." Unsatisfactory residents comprised: at the bottom, the "Drop-outs"—men who had left the field of psychiatry; the clearly "Inadequate," whom we selected, relying heavily on qualitative criterion information, from the residents who got the lowest 20 per cent of the ratings; and the "Borderline"—a small residual group of doubtful cases. (The bases of these classifications are fully explained in Appendix 8.13.)

#### RELATIONSHIPS AMONG THE VARIOUS CRITERIA

Finally, let us see how well the different approaches to evaluating competence agree. If they correlate well with each other, it supports our assumption that all are measuring what they are supposed to measure.

As we improved the technique by which we got Supervisors' and Peers' Evaluations, they agreed more and more closely, until for the last class for which the comparison can be made they correlated about .8, which is almost as high as the internal consistency of each. Moreover, both of these main criteria (evaluations of Over-all Competence) agree quite well with post-training criteria. (Again, full details on these interrelationships are given in Appendix 8.6.)

All in all, the pattern of agreements between criteria confirms our confidence that we chose the best available techniques for measuring competence of psychiatric residents. The Supervisors' Evaluations and, to a lesser extent, Peers' Evaluations were highly reliable, agreed well with other criterion information, and have a good deal of "face validity."

## THE SUCCESS OF SELECTION BY RULE OF THUMB

THE INITIAL RESEARCH PLAN WAS TO PUT INTO OPERATION AS GOOD A method of selecting residents as possible and then to see how well it worked. In this book such an approach is called *selection by rule of thumb* since it relies upon the intuitive abilities of experienced clinicians, using their favorite instruments and knowledge of what a psychiatric resident does and should be like.

Three steps made up the research method:

1. Setting up an assessment procedure patterned after the diagnostic evaluation of patients, with three psychiatric interviews and a battery of tests interpreted by a psychologist.
2. Having the four assessors who saw a candidate make independent predictions of his over-all competence as a psychiatric resident.
3. Following up the subjects, obtaining criterion evaluations and correlating them with the predictions.

Accordingly, each applicant for residency was interviewed during a two-day visit to Topeka by three experienced psychiatrists, each of whom made an independent prediction, on a ten-point scale, of how good a psychiatrist the applicant would make and a recom-

mendation (Take, Doubtful, or Reject). The psychologist gave the Rorschach, Wechsler-Bellevue, Word Association and Szondi tests, and the Strong Vocational Interest Blank; relying mainly on the first three of these, he wrote a report and made a similar recommendation and rating. The test report was freely written but usually covered personality structure, evidences of pathology and abilities for psychiatric work.

The interviewers were drawn from a large group of staff psychiatrists from The Menninger Foundation and Winter VA Hospital. They were free to develop their own styles of interviewing and to cover any areas they wished. The interviews were supposed to be an hour long, and some guidance was furnished by forms on which impressions and ratings were to be recorded.\* (See Appendix 9.1. A full discussion of the size and composition of the population of subjects is given in Chap. 4.)

Several of the interviewing psychiatrists and the psychologist who wrote the test reports made up the Admissions Committee. It did *not* operate simply by averaging all the predictive ratings a man received and applying some fixed cutting-point. It used other data available to it as well as the psychiatrists' and psychologist's reports in reaching decisions. In cases of disagreement between the assessors, some attempt was made to reconcile the differences through discussion, though practice varied as the personnel of the committee changed. The assessment and the predictive ratings were done routinely as part of the job of selection; all data were then turned over to the research workers, who never participated in administrative decisions about any of the experimental subjects.

As an aside, the interviewers agreed moderately well with one another (average intercorrelation of ratings for all 429 subjects was about .4) but slightly less well with the psychologist (correlation between the average rating by interviewers and psychologist's rating was .3). (See Appendix 9.2.) These figures mean that interviewers and testers evaluated the applicants rather differently; they tell us nothing about how effective any of the predictions were. For that, we shall have to turn to the criterion data.

\* Actually, no ratings were made on 11 per cent of the applicants (though only a few of these were not interviewed), and 21 per cent of the acceptees in our experimental group were seen by two interviewers and 5 per cent by one, instead of the usual three.

## THE VALIDITY OF THE PREDICTIONS

## COMMITTEE DECISIONS

Examining first the decisions of the Admissions Committee of the School of Psychiatry, we find a respectable degree of validity (see Fig. I-9.1). Accepted applicants remain in the field of psychiatry significantly ( $p < .001$ ) more often than those who were rejected. Moreover,

73 per cent of accepted applicants \* are diplomates of the American Board of Neurology and Psychiatry; while  
37 per cent of the 172 rejected candidates are diplomates

—again, a highly significant result ( $p < .001$ ).

To an unknown extent, some allowance must be made for the fact that rejection by the Committee affected the decision of a number of men to cease efforts to specialize in psychiatry. Some of them wrote either to the School or on our Follow-up Questionnaire that rejection by the Menninger School of Psychiatry added to their doubts about going ahead with their original career plans. Their replies to our questionnaires on reasons for changing plans were not very revealing, however. Most said either that rejection by this and other schools deterred them or they discovered their interests could be better expressed in another specialty, e.g.: ". . . I feel that I can do more good, be more effective psychologically and generally in pediatrics. I do not care to divorce myself from organic disease per se." Also, much of the better showing of the acceptees seems due to the fact that most of them went to the Menninger School (see Appendix 9.3).

## INTERVIEW AND TEST RATINGS

As could be expected (and is shown in Fig. I-9.1), the interviewers and testers performed rather similarly to the Admissions

\* Among the 238 alumni of the Menninger School of Psychiatry, 78 per cent are diplomates, but only 48 per cent of the 46 applicants who were also accepted but went elsewhere for training are diplomates. All of these figures refer only to certification in psychiatry; a number of subjects who have gone into other medical specialties are diplomates of the appropriate boards. Information was obtained by checking lists of subjects against published names of those who passed after each Board examination from 1946 through 1956.

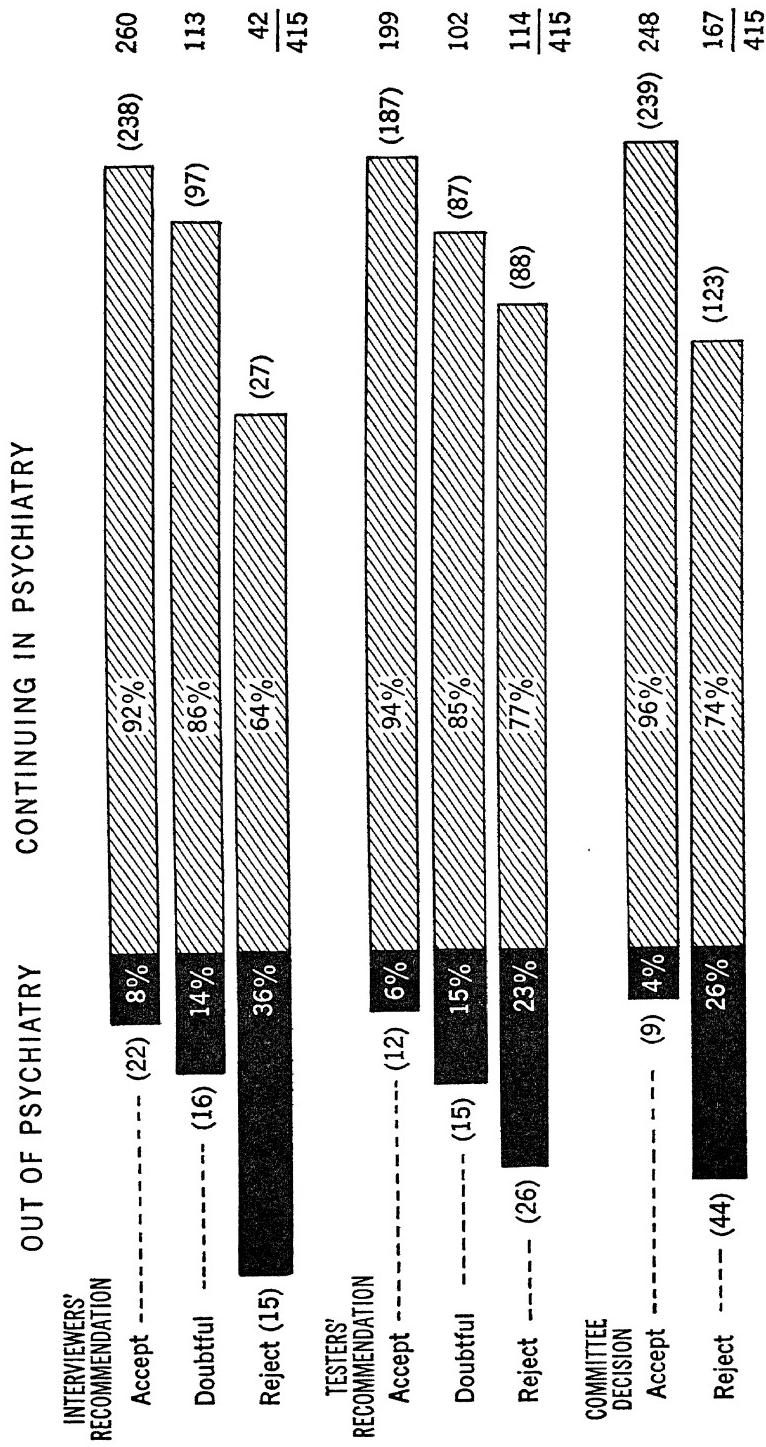


Fig. I-9.1. PROFESSIONAL STATUS OF APPLICANTS TO THE MENNINGER SCHOOL OF PSYCHIATRY AS OF JULY 1953

Committees, since the latter's decisions were based on data and recommendations they supplied. Both the testers and interviewers recommended for acceptance people who later tended to stay in the field of psychiatry, as compared to the applicants they thought should be rejected. (Comparing the proportion staying in psychiatry among those Accepted and those Rejected by testers, the difference is significant at the .001 point. The corresponding difference for the average judgment of the psychiatric interviewers is also very highly significant, at the .001 point.)

Moreover, these recommendations predict certification by the American Board of Psychiatry and Neurology also, again with very high statistical reliability ( $p < .001$ ).

- 66 per cent of applicants whom interviewers accepted were diplomates by the end of 1956;
- 52 per cent of those about whom interviewers were doubtful were diplomates;
- 37 per cent of those they *rejected* were diplomates.

The recommendations of the testers foreshadow success on this criterion even better.

- 76 per cent of applicants whom testers accepted were diplomates;
- 55 per cent of those about whom testers were doubtful were diplomates;
- 41 per cent of those they rejected were diplomates.

(More detailed figures on the validities of the numerical predictive ratings are given in Appendix 9.3; in general, they tell the same story.)

In terms of the basic follow-up criteria, then, all three major predictors had significant usefulness. But how did they differ in their relative validities? It is not easy to see at first whether testers or interviewers were superior in picking men who would stay in psychiatry. Looking again at Fig. I-9.1, if we ask how the applicants who were recommended for rejection by testers turned out as compared to those rejected by consensus of the interviewers, the interviewers' pooled judgments seem to be a better guide. It should be remarked, however, that interviewers recommended only 42 out of the 415 men for rejection—about a third as many as the psychologists wanted to reject. Perhaps the interviewers felt it appropriate to reject only clearly unsuitable men.

We can easily test this hypothesis by considering groups of equal size who were given the lowest ratings by testers and by interviewers. If we take the 15 per cent who received lowest ratings from the psychologists, we find that 38 per cent of them turned out to be of borderline adequacy or worse or to have left psychiatry, whereas only 21 per cent of the interviewers' lowest 15 per cent turned out to be poor in those ways. Thus, the tests gave rise to recommendations significantly more often correct than those made by the consensus of three interviewers if we equate for differences in the numbers recommended for rejection. (Chi-square for this difference is significant at the .05 level.) The difference in the charts in Fig. I-9.1 can thus be attributed almost entirely to differences in criteria for acceptance.

If the interviewers' recommendations alone had been the basis for acceptance, and even if everyone about whom they felt doubtful had been rejected, more men would have been admitted than the School could have handled, including twenty-two who eventually left the field. If the psychologists' recommendations after testing had been used, too few men would have been admitted, including however only twelve Drop-outs. The committee was able to learn something from both their more optimistic and pessimistic informants, accepting among the number of men they wanted only nine who were unable to continue psychiatric training or decided to leave it.

Using the post-training criteria of "staying in the field" and of Board certification, *the decisions of the committee tended then to be somewhat more often valid than either testers' or interviewers' recommendations*. These statistics do not tell the entire story about the efficiency of the Admissions Committee's judgments, however. Pooling all available information on a case, the Committee could give differential weight to the opinions of various assessors, bring in information that emerged only in letters of recommendation or the like, and give special weight to information elicited by only one interviewer. For example, one candidate managed to make a pretty good impression on two interviewers and the tester, but in his third interview broke down and confessed to being an overt homosexual and to having made a suicidal attempt. The interviewer persuaded the man to withdraw his application and seek personal treatment. If the ratings had simply been averaged, he might well have been admitted.

The Admissions Committee thus succeeded in excluding a number of preschizophrenic and borderline schizophrenic applicants, extreme sexual deviants and other severely disturbed persons. Moreover,

we know that the Committee weeded out those with low intelligence as well as many who lacked strong interest in psychiatry. We have no way of knowing how efficiently people who would have been inadequate on other grounds were eliminated. Nevertheless, considering the types of liabilities among those rejected, we can see that the range of talent or suitability for psychiatry among the residents in the School was considerably smaller than in the entire group of applicants.

Nevertheless, we can learn a good deal more about the ability of interviewers and testers to assess psychiatric suitability by using in-training criteria, which are of course available only for acceptees who were trained in Topeka. When we correlate the predictive ratings with the *supervisors' evaluations*, we find that both interviewers and testers were able to predict significantly better than chance. The validity coefficient for the *average* of interviewers' ratings is .24; for the psychologist's test ratings it is .27 (see Appendix 9.4). The difference between these two correlations is not significant.

Validity figures of this size are high enough to make it worth while to use the predictors if nothing better is available, but they mean that neither the pooled judgment of psychiatrists nor the laboriously-arrived-at ratings of psychologists could be relied on heavily to predict just how satisfactory a resident an applicant would make. In evaluating these correlations, however, remember that the restricted range of talent in the accepted group makes high validities statistically unlikely.

Both of these predictors were correlated with the Supervisors' Evaluations for each of the six classes of residents separately and with the Peers' Evaluations for the classes on which they were available. These results (see Appendix 9.5) show primarily that there was a good deal of fluctuation in the validity figures from class to class, probably due to differences in the samples of applicants and in the persons making the predictions. There was no clear-cut trend for superiority of either interviews or tests as predictive data; validity coefficients for both varied a great deal, now one being higher, now the other.

Some results from miscellaneous criteria may be briefly summarized (they are given in full in Appendix 9.6). Correlations with *academic examinations* were generally low and insignificant, though men who were thought promising by the interviewers tended to do somewhat better on the exams than those who were rated low by

interviewers. The *follow-up criterion* of success early in practice could not be correlated with predictive ratings, but it showed that applicants accepted by the Admissions Committee tended to make more rapid professional progress than those rejected. Finally, the occurrence of psychiatric casualties may be considered a criterion. The Admissions Committees were specially concerned to eliminate persons who might develop severe personality disturbances. We therefore examined the test and interview predictions that had been made on 10 known psychiatric casualties. It should be noted that all had been accepted by the Admissions Committee; there may have been many more among the rejectees, but we had no way of getting such information on them. Neither testers nor interviewers gave any of these subjects unusually low ratings—the test ratings ranged from 6.8 (mean of 6.8) and the interviews' ratings from 6.0–9.1 (mean of 7.6). (It was not at all unusual for accepted cases to be given test ratings as low as 6.) If the testers' recommendations had been followed, 3 would have been eliminated, but at the cost of many adequate residents. It does not seem, therefore, that simply raising the cutting point of acceptability would definitely have prevented these mistakes.

All of these casualties occurred in the earliest classes of the School. It is possible that such precariously adjusted people did not apply for later classes, but it seems more likely that the School's Admissions Committee learned from unhappy experience how to weed them out.

#### THE ACCURACY OF INDIVIDUAL PREDICTORS

Since there was a sizable group of psychiatrists and psychologists working with the Admissions Committees, one should expect—just on the basis of random variations—that their predictions would differ somewhat in validity. Their validities do vary, but only to about this expected degree, and cannot convincingly be related to the general clinical skill of the predictor. Predictive ratings by the most successful interviewer, for example, correlated .27 with Over-all Competence, only slightly (and not at all significantly) better than the average rating for all interviewers. Though most of the interviewers and one of the testers were unable to do significantly better than chance in making finely graded predictions (as measured by correlation coefficients and percentages of extreme errors; see Appendix 9.7 and Table 9.10), almost all did significantly better than chance in

making the basic distinction between satisfactory and unsatisfactory applicants.

#### ONE, TWO, OR THREE INTERVIEWERS

Thus it appears that with only one interviewer, one should not put a great deal of faith in his judgments beyond the decision to accept or reject, and even that cannot be taken for granted until proved to be better than tossing coins. In correlational terms, the average validity of predictions by a single psychiatric interviewer is .19. When two interviewers' ratings are averaged, the validity of the result goes up to .24, and for an average of three interviewers' judgments it becomes .27.\* The remarkable and comforting fact is that a modest amount of truth seems to emerge from a pooling of error. One interviewer's biases must often have balanced another's pretty well.

Such were the results of Design I: They were good enough to warrant considerable satisfaction but left plenty of room for improvement. The findings especially endorse the method of having various sources of information (including the ratings and recommendations of psychiatrists and a psychologist who had interviewed and tested the applicant) brought together into an integrated picture by an Admissions Committee, which could then make the actual decision to accept or reject. This was, however, an expensive method, and nothing in the results indicated how it might be streamlined and made more efficient. There was a hint that psychological tests were especially promising but not so much so that interviewing could safely be abandoned. Clearly, if progress was to be made in learning how to do the job better, a different research design was needed.

\* Taking the group of 149 residents who were interviewed by three psychiatrists, we randomly chose a single interviewer's rating for each case and, again, randomly divided the remaining pairs of ratings into two more series. These were then combined three different ways and averaged. The figures (.19 and .24) represent the average of the three validity coefficients for the series of single ratings and of paired ratings. The difference between correlations for one and for three interviewers approaches significance ( $p < .10$ , one-tailed test). (See also Appendix 9.8.)

PLANS FOR A NEW  
EXPERIMENTAL DESIGN

SHORTLY AFTER WE BEGAN WORK ON THIS PROJECT, THE DECISION WAS made to work intensively, rather than extensively. Instead of trying out hundreds of ratings, test scores, interviewers' impressions, miniature life-situation results, etc.—as predictors of professional success—it seemed that we would do better to concentrate on a few techniques and set up our design in such a way as to gain maximal understanding of all the processes involved. If a promising-looking technique turned out to be valueless, we wanted to know why. If it should be impossible to predict total performance, perhaps the fact that certain aspects of it were actually well predicted was concealed by miserable performances in predicting others. Or, if there should be encouraging successes with a particular test, we wanted to be able to say *how* they were achieved.

To do a good job of selecting psychiatric residents, we must ask, first of all: *What does a psychiatric resident do?* This means a functional job analysis of the work done by a resident (many elements of which carry over to jobs filled by trained psychiatrists). Next, we must know *what to look for*: What aspects of the person—his personality, abilities, and history—are susceptible of being assessed in a short time and are relevant to his later performance in these psychiatric functions? *How should we look for it*—what techniques and

procedures of assessment would be most fruitful in enabling us to evaluate the relevant aspects of a person? Then: *How can we recognize it when we see it*—what are the most efficient (reliable and valid) methods of analyzing the data yielded by assessment methods so that we can best distinguish between men who can become good residents (and psychiatrists) and those who cannot? Finally, will the results obtained hold up with a new sample of subjects? Each of these questions presupposes an answer to the previous one, and we were convinced that we should have to provide some kind of answer to all before the project was over.

We decided on a five-fold plan: (1) To learn the duties of a psychiatric resident, we consulted with a number of psychiatrists from Topeka, and elsewhere, and worked out a set of definitions for the major work-functions and for Over-all Competence as a psychiatric resident. (2) To understand what to look for, we sought the opinions of experts: people engaged in the job of selection who had definite ideas about what was, and was not, relevant to psychiatric work. A further step was equally essential: to make a careful study of men who actually did either well or poorly in the residency program. This step, which we came to call the "Small Sample Study," played a crucial part in also providing answers to other questions. (3) In studying intensively our best and poorest residents, with some others in intermediate ranges of proficiency, we tried out a large number of assessment techniques (tests, interviews, and documentary materials). (4) With each procedure, we concentrated on trying to find ways in which it could be used to differentiate the better from the poorer residents—first, in a somewhat impressionistic way and, later, through the slow and laborious writing of manuals for the analysis of each technique. Whenever we found a test response or an aspect of an interview discussion that good residents handled differently than poor ones, we asked ourselves: What personological concept do we need in order to conceptualize and generalize this finding? The manuals, therefore, were written as handbooks to enable anyone to extract a group of personality variables from the particular assessment techniques involved. (5) Taking our best hypotheses, methods, and findings, we tried them out in a predictive study, in which we forecast the performances in the School of a group of new subjects, about whom we knew only our assessment data.

In the job that we were undertaking we had to hew out of raw materials each plank that went into the finished structure. For there

were no satisfactory job analyses of psychiatric work; there was no assured knowledge of what a good, or poor, psychiatrist was like in any detail; there was little more than custom and prejudice to guide the choice of assessment techniques; and no systematic methods were available for reliably and validly analyzing the results of psychological tests and interviews in terms of the personological variables we needed to use. Looking back upon the size of the task we undertook, we are astonished at our brashness in taking on so much with such slender resources.

**JOB ANALYSIS**—Although we began work on this aspect of the research in our first month, we only gradually realized just how seriously we should have to take it (see Appendix 10.1). By the time we got around to writing definitions of the work functions, we had had a large number of evaluative criterion interviews, in which people who were well acquainted with a resident's work discussed it with us in detail. From this experience, from the suggestions of our psychiatric consultants, and from materials made available to us by Drs. Kelly and Fiske of the Michigan project, we settled on the following functions and definitions:

*Psychotherapy*: Conduct various types of individual psychotherapy, flexibly adapting therapeutic approach to the needs of the particular patient; or work with a narrower range of methods, restricted to the kinds of patients to whom they are appropriate.

*Management of patients*: Work with patients in briefer contacts than those of formalized psychotherapy. Make proper decisions regarding the day-to-day treatment of the patient and attitudes to be adopted toward him by hospital personnel. Handle appropriately the patient's complaints and requests. Become familiar with all possibilities of treatment, prescribe and integrate a treatment program appropriate to the patient's problems, resources, and total situation; manipulate the total therapeutic environment of a patient. Establish good working relationships with the patient's family.

*Diagnosis*: Elicit from the patient information required for mental status examinations and case histories; ascertain and evaluate attitudes and incidents of psychological significance in the patient; observe and interpret significant aspects of speech and behavior.

Relate general concepts and descriptive (nosological) categories to concrete individual syndromes, flexibly yet accurately; recognize dynamics underlying the patient's behavior and life history. Synthesize clinical findings to arrive at an integrated picture of personality development, structure and function, and the nature of decompensation.

*Administration:* Plan and develop over-all psychiatric programs; or maintain an efficiently functioning milieu in which the patient can recover. Know and utilize the special contributions of all members of the psychiatric team, elicit their co-operation and maintain their morale; delegate responsibility appropriately.

*Over-all Competence:* Carry out effectively any useful role in psychiatry, which may include one or more of the first four functions above, or others (such as research).

Early in 1949, we made a more thorough study of the evaluation interviews we had had with sixteen of our most articulate raters on the first four classes of residents. From the transcripts, we extracted all concrete references to particular instances of good and poor psychiatric work, then grouped these statements into twenty-two general categories. These we called "specific aspects of work." (They are listed in Appendix 10.2; a selection from this list became our "minor [predictive] variables.") In some ways, this was a rather crude approximation to the method of critical incidents that has been worked out by Flanagan and his associates (1954) for job analyses. These two sets of variables ("major" and "minor") served as the job-analytic basis for the final predictive study with which the project was completed. They do not completely cover the work of the resident, but we considered our task to be getting together a reasonably, not exhaustively, comprehensive list of work functions, of a kind and in such numbers that we might hope to be able to predict them.

**SURVEY OF OPINIONS ON THE PERSONAL QUALITIES OF GOOD AND POOR RESIDENTS**—The results of this phase of the work were so extensive and so interesting in their own right that a full account of the survey is presented separately (Chap. 15).

**SMALL SAMPLE STUDY: CHOOSING ASSESSMENT TECHNIQUES**—So far we had dealt only with hunches, hypotheses, generalizations, and distillations of experience in psychoanalytic training. No matter how well grounded all this might be theoretically, we knew that there was no substitute for a clear look at facts: a study of residents who had turned out well and residents who had performed badly. The evaluations we began gathering in October 1947 formed a base on which we chose sixteen "Highs" and thirteen "Lows," plus three "Middles." The spring of 1948 was devoted very largely to studying these men. (For the methods by which they were chosen see Appendix 10.3.)

Our first purpose in working with these extreme subjects was

to try out a variety of tests, interview approaches, and other sources of information that might be useful for selection. We knew there was some danger in it since the men were in a situation very different from the one of being assessed for possible training. Yet there seemed to be no alternative. It would take too many years to try out all promising procedures by giving them to candidates when they came to apply and then, later, correlating test scores with performance in the School. It would be necessary to take the chance, we thought, and surely a method that clearly distinguished Highs from Lows at the end of training might be useful for selection.

By the beginning of 1948 we had gathered enough ideas from various authorities to form a list of over thirty personal qualities thought to be necessary for psychiatric work. It was very helpful to us, in making a rough assay of the methods being used in the School's regular admissions procedure, to see what areas of personality were being neglected or not very securely measured. To span the gaps that were discovered in this way, we surveyed available methods, settling upon several published, and unpublished, tests, a systematic interview taking several hours, an autobiography, and a few tests we devised ourselves. In making these selections, we were guided by the principles of *feasibility* and *adequacy* to the task. Some promising ideas were discarded because of practical and financial objections to their use; others because of the superficiality of their approach to the personality. In the former group belonged role-playing and other situational tests; in the latter, most of the published "personality tests" that are based on self-report. We stressed projective tests and interviews, though we did try out some standardized, and unstandardized, paper-and-pencil questionnaires.

Counting the tests that the subjects had taken when they applied to the School, by the time we were through they had been given sixteen different tests and two interviews. Since most of these tests died a rather quick death, there seems little point in devoting much space to them. (A description of each may be found in Appendix 10.4.)

Six procedures were selected as worthy of further study on the basis of their ability to discriminate Highs from Lows, as well as their practicability for a selection program. These were: Credentials, the structured Interview, TAT (Thematic Apperception Test), Self-Interpretation of the TAT, Rorschach Test, and Picture Reaction Test (PRT). In addition, the Strong Vocational Interest Blank,

Wechsler-Bellevue, and Word Association tests also seemed promising, though we never actually constructed manuals for them. At any rate, we had now chosen a number of ways to assess candidates, intending in subsequent work to shorten some of them, eliminate overlap, and discover combinations of less than the total number that might provide enough information to warrant our recommending them for routine use in selection.

**SMALL-SAMPLE STUDY: DEVELOPING MANUALS TO GUIDE CLINICAL JUDGMENT**—We had now collected some plausible hypotheses about aspects of personality that were relevant to a man's functioning as a psychiatrist. We had studied small groups of superior and unsatisfactory residents and could propose a battery of assessment methods for further study on an empirical basis other than "clinical experience." Before proceeding to examine the business of preparing manuals, we should note one other important point about the Small Sample Study.

At the time we started developing tests to try out on the Small Sample, we had put together, with the help of authorities, a list of personological concepts which appeared to be quite usable. The next step seemed to be to consider each of these thirty concepts carefully, attempting clarification of them by means of theoretical and operational definitions. Consequently, the five of us (Rapaport, Holt, Luborsky, Morrow and Ramana) set to work writing definitions. In a long series of meetings, we discussed each other's definitions and tried to arrive at some consensus. Although we all approached the problem with Freudian psychoanalytic orientations and did reach mutually acceptable definitions for a few concepts, the process was entirely too difficult and time-consuming to be carried to conclusion.

We turned instead to our data on the personalities of the Small Sample Highs and Lows, making use of any concept that seemed to capture best the personality differences we could see.

As each worked out his manuals, then, he tried to put together whatever characteristic the two groups had in common, letting consistencies in the data imply use of either new concepts or those that had been suggested by our study of experts' opinions. As it happened, both of these possibilities were realized, and there was an encouraging degree of convergence in the various lists. Of course, it was not to be expected that differences in observational ability or perceptual sensitivity, which showed up very prominently in differentiating the performances of the Highs and Lows on the Picture

Reaction Test, would emerge in the Interview. Nevertheless, there was nothing in the Interview to suggest that such a concept might not be a useful one.

The process of developing a list of variables satisfactory to all three of the main experimenters was a long one, going hand-in-hand with the development of the manuals. Here is the final list of personality variables with which we worked throughout the rest of the study. Each concept is followed by a parenthetical H or L, indicating that the variable was defined in terms of the high (H) end of the scale or the low (L) undesirable end. The names of the latter variables are preceded by "freedom from," indicating that high ratings were always given to the desirable end of each continuum.

#### MOTIVATION FOR PSYCHIATRIC WORK

1. Freedom from Material Preoccupations (L) : Money-mindedness; a need for a comfortable, lucrative, and easy kind of medical practice.
2. Freedom from Status-Mindedness (L): Need for medical prestige, social status, respect, or fame as the principal rewards of psychiatric work.
3. Need to Help (H) : Kindliness, the drive to help others and relieve suffering.
4. Psychological-Mindedness (H): An interest in, orientation toward, and/or ability to grasp the psychological dimension of experience; particularly, an understanding of behavior in terms of its motivation.
5. Psychological Curiosity (H): A drive to find out and understand what makes other human beings act the way they do.
6. Internalized Interest in Work (H): The capacity to become interested in working and achieving for the sake of the intrinsic rewards of the process, rather than the extrinsic gains of money, notoriety, etc., that it may bring.
7. Internalized Interest in Psychiatry (H): Similar to No. 6, with specific focus on psychiatric work—a need or capacity to enjoy it for its own sake.

#### SELF-ORIENTATION

8. Self-Confidence (H) : A basic (*i.e.*, not a superficial protest) trust and faith in oneself; the capacity to maintain a healthy degree of self-esteem with a relative lack of external support.
9. Self-Objectivity (H): Insight into oneself; a freedom from blind-spots, exaggeratedly self-glorifying or self-debasing attitudes.
10. Sense of Humor (H) : A capacity to enjoy the ridiculous or comic aspects of situations, including the ability to relax and have fun when it is appropriate to do so, and to be witty in a non-defensive way.

11. Capacity for Personality Growth (H): A positive orientation toward change and development; the ability to learn not only intellectually but emotionally.

**ORIENTATION TOWARD OTHERS**

12. Objectivity Toward Authority Figures (H): An ability to react appropriately and realistically toward one's superordinates (or parental surrogates).
13. Objectivity Toward Dependent Figures (H): The capacity to react to patients (and to children or others who may symbolize patients) in a realistic way, again without crippling transference or counter-transference manifestations.
14. Empathy (H): The ability to take the role of the other in imagination, to understand or to "resonate to" another's feelings and viewpoints.
15. Freedom from Ethnocentric Prejudice (L): Rejective, usually stereotyped and/or moralistic attitudes toward members of ethnic minority groups.
16. Consciousness of Social Injustice (H): A generally humane and humanitarian outlook toward social issues; identification with the underdog.

**SOCIAL ADJUSTMENT**

17. Mature Heterosexual Adjustment (H): Capacity to establish mutually loving, genital (in the psychoanalytic sense) relationships with a woman.
18. Social Adjustment with Co-Workers (H): Congenial and mutually helpful relationships with colleagues and other members of the psychiatric team.
19. Feelings of Security (H): An emotional attitude of being loved and supported by others, as opposed to feeling alone and rejected in a cold, bleak world.

**HANDLING OF EMOTIONS**

20. Adequate Emotional Control (H): This variable, and the next, were defined so that the highest point on the scale they referred to represents an optimal balance of freedom and control of impulse and affect. The low end of the scale may denote either *too much* control (inhibition, isolation of affect, defensive withdrawal, etc.) or *too little* control (impulsiveness, emotional vulnerability and lability, etc.).
21. Adequate Handling of Hostility (H): The capacity to express anger, irritation, etc., at the right times, to the right persons, and in the right amount, as opposed to undercontrol or overcontrol of the kinds mentioned just above.
22. Emotional Appropriateness (H): Like No. 21, No. 22 is also a specific aspect of No. 20 abstracted from it: finely differentiated affective responsiveness that is sensitively attuned to the demands of the total situation.

INTELLECTUAL EFFECTIVENESS

23. Clarity of Thought (H): The ability to deal effectively with ideas, particularly psychiatric ones, without confusion, vagueness, or muddle-headedness.
24. Freedom from Stereotypy of Thought (L): Conventional, unoriginal manipulation of clichés, as opposed to original and imaginative thinking.
25. Perceptual Sensitivity (H): The ability to be observant, to pick up subtle, yet psychiatrically important, aspects of the perceptual field.
26. Cultural Wealth (H): Richness of cultural background, possession of good taste and genuine interest in the arts and humanities.
27. Quality of Verbalization (H): The ability to use language sensitively and effectively.

HONESTY

28. Well *vs.* Poorly Internalized Ethical Standards (H, L): A state of having developed basic values and standards about conduct that one believes in deeply, in contrast with being ruled by fear of shame, discovery or punishment from the outside, rather than by an inner feeling of what is right and wrong.
29. Genuineness *vs.* Façade (H, L): The capacity to be oneself, naturally and without pretense, as opposed to character defenses involving falseness, artificiality, or attempts to be, or seem to be, something that one is not.

PATHOLOGICAL TENDENCIES

30. Freedom from Overobsessiveness (L): At the lower extreme, a state in which effective action is paralyzed by doubt, ritual, and rumination.
31. Freedom from Projection (L): More than normal reliance on the psychoanalytically defined defense mechanism of attributing unacceptable tendencies in oneself to others, which, at the extreme, entails paranoid suspiciousness, querulousness, grandiosity, etc.
32. Freedom from Schizoid Disorganization of Thought and Affect (L): At the lower extreme, this variable would indicate a fully developed schizophrenic process. We operated on the assumption of a continuum of autistic, or schizoid, manifestations of thinking disorder (loose or concrete conceptual thinking, disrupted concentration, peculiar use of language, etc.) and disorganization of affect (blandness, inappropriateness, disharmony, etc.).

After we had finished working through a number of cases together, we found we had no difficulty in using these concepts in the same way, and so we never felt the necessity to adopt rigorous and "official" definitions. Our concepts were arrived at empirically; they came from many different frames of reference, existed on different

levels of abstraction, and partly overlapped with each other. In practice they proved to be quite workable and helpful. With their aid, we achieved some uniformity in the terms used in the various manuals. They represented aspects of personality that could be directly inferred from assessment data and serve as bridges, or intervening variables, leading to predictions of specific kinds of psychiatric work behavior.

The next step was a crucially important one: providing guidelines to the interpretation of data gathered by our assessment procedures so that psychiatric performance could be predicted most accurately. Moreover, predictions had to be *reliable*, in the sense that the method of using a test or interview for this purpose would not remain the private knowledge of any one man. We knew that "the interview" *per se* is not something that can reasonably be said to have any particular degree of validity or usefulness. An interview is valuable, or worthless, depending on the man who uses it and the specific way in which he conducts it and interprets the findings. The same thing goes for diagnostic psychological tests. Not only is this true of such obviously projective and qualitative instruments as the Rorschach and TAT, but even "objective," self-administering tests, such as the Strong and Minnesota Multiphasic, vary enormously in their usefulness, according to the skill and experience of the clinician who interprets the quantitative profile of scores.

Our aim and hope, then, were to guide the judgment of the clinician who would use our tests and interviews, by means of handbooks, or manuals, which would tell him what to look for and how to interpret it. Each of us (Holt, Luborsky, and Morrow) undertook the construction of two manuals. Dr. Walter Kass furnished a manual for the Rorschach, which was based on our Small Sample and developed in the same way as the others. We were not able to work up manuals for the Wechsler-Bellevue and the Word Association tests, which were given a subordinate place in the experimental design.

The first stage in the process of writing a manual was detailed and repeated study of the protocols from the Small Sample subjects. Some of the differences found between Highs and Lows were relatively specific (*e.g.*, only Lows saw volcanic eruptions in the Rorschach blots); others had to be sought on a higher level of generalization (*e.g.*, Highs excelled in the degree to which characters in their TAT stories were made vivid and psychologically credible).

All of the varying kinds of differentia we speak of as "cues." We had to think of the cues, in turn, as indicating aspects of personality, so that they could be used flexibly in making several kinds of predictions. Our clinical experience, theoretical background, and knowledge of relevant specific findings in the literature guided us in deciding what variable was operating to bring about the appearance of any differentiating cue (see Appendix 10.5).

The time-consuming job of identifying the cues to high or low performance, and their significance in terms of personality variables, was begun in the summer of 1948 and took more than eighteen months! (Appendix 10.6 includes sample pages from each manual, as well as an outline of the Interview manual; see also Appendix 11.1.) The complete manuals are quite voluminous, due to the quotations of illustrative material for each cue (examples to which it applies definitely, examples to which it applies partially or doubtfully, and examples of similar material having different significance from the cue in question). The manuals contain, in addition to this basic set of guide lines, indications of relative weights to be given each cue, and (in some cases) a few "principles of pattern analysis"—*i.e.*, how to evaluate certain combinations of cues that have other than summative significance in arriving at the final predictions.

THE "WARM-UP"—Before the Predictive Study proper could start, we had to learn to use each other's manuals. Though we had been working in fairly close collaboration, we knew the manuals of one another only to the extent of critical readings and discussions of problems in their construction. We therefore spent a couple of weeks trying out the new manuals with cases we were not going to use in the study proper. There were plenty of protocols of each procedure to fit this description: all tests and interviews we had given to applicants who were rejected or who decided not to follow up their acceptance.

The procedure differed somewhat from manual to manual; some of them we went over more thoroughly than others. Since "Judge II" was not going to score the PRT manual, or "Judge I" the Interview, there seemed no point in their mastering the scoring processes involved. When a judge was to use a manual in the study, however, we tried to have him score ten cases and check his scoring against that of the manual's author, discussing differences and reaching agreement. We were enough impressed with the difficulties of achieving

reliable scoring to consider check-scoring a sample case every so often during the series of actual Predictive Study cases. Indeed, it would have been most desirable, from the standpoint of attaining good reliability, to have spent more time in this kind of training at the start, and throughout, to make certain that we did not diverge in our techniques; but we had already delayed the beginning far too long.

**THE PREDICTIVE STUDY: EXPERIMENTAL DESIGN**—The Predictive Study was to be a cross-validation—the application to a check sample of the procedures and concepts that worked with the original Small Sample, to see how well they would hold up. It was also to be our final effort. Accordingly, we built into it as much as we could, hoping to find answers to many questions about selection and related issues. The result is a complicated experimental design.

In the attempt to get data on both reliability and validity of the manuals, each had to be applied by at least two judges. The relative validity of individual procedures, and batteries composed of these procedures in their principal combinations, was to be tested; also we wanted to find out whether there was an optimal amount of information short of a maximum. Consequently, each judge was to start every case with a single source of data, make his judgments, add another procedure, make the same judgments again, and continue in this way until he had been through the complete file on the subject. At each of the first stages of analysis, the judge first applied a manual rigorously and then made free ratings based on his total clinical impression to that point. These aims then determined the general design of the way data were fed to each judge and the general types of responses expected from him.

The specific judgments required at each stage of analysis were further determined by a few other issues. What is the optimal level of generalization toward which to aim in predicting behavior? Should one try to stick to concrete particulars of reaction in specified situations, or should one aim at very general summary statements about the whole trend of a man's behavior? To answer these questions we had to make predictions on four levels of specificity at each stage of analysis. In addition, we had a list of personality variables, all supposedly relevant to a man's psychiatric functioning and all in need of cross-validation. Therefore, we also had each judge rate all thirty-two personological variables at each stage. And so that the halo effect

of emotional bias for, or against, a subject might be studied and controlled, every judge had to round out each set of predictive ratings by rating his personal feeling of like or dislike for the subject.

Finally, there was the over-all objective of broadening as much as possible the scope of our understanding of the predictive process, its successes and its errors. Many aspects of the design already mentioned helped in this respect, but a final and most burdensome provision was also added. The judges were to record, to the best of their ability, the bases of each judgment made about the applicant—the raw data in question, the interpretation made and, if possible, the principle underlying that interpretation.

The rich mass of data, ratings, and interpretations on each subject would thus make an excellent base from which further studies could be made in the future. We had in mind, for example, that someday, when the follow-up data were in, groups of subjects who differed in the types of practice they had set up could be studied intensively for features of personality they might have in common. To facilitate such studies and to round out each clinical case study, the judge was asked to write a brief integrative and synthetic sketch of the personality after making his personality ratings and before making his predictive ratings—again at every stage of analysis.

#### SUBJECTS

The original plan was to get all procedures worked out so that we could apply this elaborate analysis to every applicant for two years. In this way, we hoped to get about ninety accepted subjects, whom we should be able to follow through the School, plus enough rejectees who would drop out of psychiatry, to bring the grand total to one hundred.

As the planning for the final study went ahead, with more judgments and more work at each stage of analysis constantly being added, it became clear that we probably could not analyze the data as fast as we collected them. It was clear, too, that the manuals were not going to be ready on time. Indeed, most of them were hardly further along than first draft when it became necessary to start testing the first subjects of the final group. Though we did not realize it, we were to finish the testing of all subjects before starting to analyze any of the data! Consequently, we were spared the waste of elaborately analyzing subjects about whom we should have had very little follow-up.

Since we needed, in addition to the regular test battery and credentials that were being routinely obtained, an additional application form called "Supplementary Face Sheet," the recorded systematic Interview, the TAT, Self-Interpretation, and Picture Reaction Test,\* we had to schedule an additional day for each applicant. They were not told which procedures they underwent were experimental and which were operational; indeed, they did not suspect that any would not be used in their selection. Thus, except for the added strain of a three-day, instead of a two-day, assessment, the data were gathered just as they would be in actual practice.

A note here about time sequence. The Small Sample data were gathered during the spring of 1948, and on the analysis of these data the construction of manuals was based. During that summer, the essentials of the final Predictive Study were conceived, and in September 1948 the first subjects in Class V were assessed by the Admissions Committee. We continued to test and interview subjects until December 31, 1949, by which time Class V had been in the School six months. Yet it was not until late in May 1950 that we were able to begin the long series of predictive judgments that lay ahead of us. During the interval—besides collecting raw data, continuing the development of criterion measures, and finishing the manuals—we had carried out preliminary cross-validations of all manuals except those for the Interview and the PRT.

In May 1949 the war broke out in Korea. This event had a marked effect on our plans. There was a noticeable dropping off in numbers of applicants during the balance of 1949 and an increase in the number of men who were accepted but never came. In the end, we had only sixty-four acceptees, plus four Drop-Outs.

#### FIRST CROSS-VALIDATION OF MANUALS

Appendix 10.7 presents in detail the results of the first predictive trial of the manuals on new samples. The TAT manuals seemed particularly promising: the Content manual's cue-sum correlated .31 with Supervisors' and .37 with Peers' Evaluations of Over-all Competence, whereas the total cue-sum from the Content and Formal manuals correlated .38 and .43 with these criteria, respectively. The Self-Interpretation of the TAT looked promising as a predictor of

\* For much of the first year, until it became clear that we could not develop an adequate manual for it, the Picture-Story Matching Test was given each applicant to Class V, along with the PRT.

Psychotherapeutic Competence; the cue-sum for this procedure correlated .39 with Supervisors' Evaluations.

The Credentials manual gave a validity coefficient of .27 on first cross-validation and seemed especially promising as a screening device to eliminate a few most unsuited candidates. Like the validities for the other manuals, this was as good as the general level of Design I validities. Only the Rorschach manual did not give encouraging results on first cross-validation, but it was extensively revised and looked as if it would work better on the next sample. Other manuals also were put through item analyses to eliminate some invalid cues, and some were rewritten to increase their discriminating power.

#### ASSESSMENT PROCEDURES USED IN THE PREDICTIVE STUDY

To a large extent, procedures carried over from the Small Sample Study were kept unchanged, but a few modifications are recorded in the following description. (The first three tests were given by psychologists working with the Admissions Committee, as part of the standard battery actually used in selection.)

*Rorschach Test.* This familiar "ink-blot test," one of the earliest of the projective techniques, provides a rich sample of perceptual-associative behavior. The technique by which it was administered and scored was generally the one described in Volume II of *Diagnostic Psychological Testing* (Rapaport *et al.*, 1946), though one psychologist administered and scored some of the Rorschachs according to the Beck system (1944).

*Word Association Test.* Known since the early experiments of Jung, the technique involves reading a list of words to a subject and requiring him to reply as quickly as possible with the first word that comes to mind. The sixty-word list and the method of administering and interpreting it are also described by Rapaport.

*Wechsler-Bellevue Scale.* This was, at the time, the most widely used test of adult intelligence. It provides (besides a total IQ) measures of Verbal and Performance abilities and consists of eleven subtests, each of which taps a somewhat different psychological function (see Rapaport *et al.*, 1945). It was scored according to the first edition of Wechsler's *Measurement of Adult Intelligence* (1939).

*Interview.* The Interview was electrically recorded, and the interviewer used a systematic guide while trying to preserve as much flexibility as possible in following the natural development of the conversation. A number of topics covered in the Small Sample interview were omitted, having not proved particularly fruitful, but the Interview still took over an hour in many cases. Applicants

reported they found the experience not markedly different from being interviewed by the Admissions Committee's psychiatrists. The Interview manual listed eighty cues, which were scored on special sheets mimeographed for the purpose. These cues were, by the nature of the method, less atomistic and more judgmental than the ones in most of the other manuals.

*Thematic Apperception Test (TAT).* Henry A. Murray's test is, like the Rorschach, one of the clinical psychologist's main standbys as a source of psychodiagnostic data. The stories told by the subject, stimulated by the test pictures, contain many projections and other expressions of his personality (*cf.* Holt, 1951). Pictures numbered 1, 3BM, 4, 15, 6BM, 7BM, 8, 13MF, 11, 16, and 12M were given, in that order. Instead of the usual technique of administration, the following modified procedure was used to save time. The examiner presented the first four pictures in the usual way. As he handed the subject the first picture, he said "Make up a dramatic story, including a description of the situation pictured, what might have led up to it and what the outcome might be. Also develop your characters by including not only their actions but their thoughts and feelings." The experimenter then took down the subject's story and all extraneous remarks, questions, etc., on a typewriter. Questions were handled by repeating or paraphrasing part of the original instructions. After the first story had been told, if some part of the story (such as antecedent events) had been omitted, the examiner asked for it and told the subject whether his story was too long, too short, or just about the right length.

By the end of the fourth story, the subject almost invariably had grasped what was wanted of him. At this point, he was given the remaining seven pictures and told: "Now go ahead and write out the rest of your stories yourself. You have the hang of it; just continue the way you have been doing. Don't turn over any of the remaining pictures until you have finished writing your story for the one you have face up. You will find some special instructions as you go along; be sure to follow them." Pinned to card 11 was a note telling the subject to write a fairy story for that picture; a note on card 16 (blank) instructed the subject to imagine and describe a scene and write a story about it; and at the end were written instructions for the next test.

*The Self-Interpretation of the TAT.* After completing his TAT stories, the subject was asked to give his own interpretation of four of them (stories told to pictures 6BM, 7BM, 13MF, and 16). The instructions told the subject to read over these stories and then to tell what he thought they revealed about the nature of his personality. Thus he was urged to take the role of diagnostician—scrutinizing psychological data, which happened to come from himself, requiring self-objectivity for their proper in-

terpretation. The productions ranged from a meager and empty sentence or two, reiterating the content of the stories, to a written page of insightful comment on aspects of personal history, present personality traits and conflicts, interests, and ideologies.

Insight might be shown, for example, in the extent to which one's own motivations were seen as determining the stories, rather than chance "external causes." An example of an "internal" explanation is: "This tells something of my need for dependence which grew out of a certain kind of relationship I had with my parents." It is opposed to such "explanations" as "This was determined by a movie I saw some years ago," or ". . . this is the way people act when they are put in this situation in the picture."

Furthermore, the Self-Interpretation procedure increases the diagnostic usefulness of the TAT by providing: (1) additional associative material, (2) clearer delineation of defenses, and (3) clues to the depth of repression (from the degree to which material is acceptable to the ego).

On the whole, however, less information was obtained from this self-administering form of the Self-Interpretation than had come from the interview-like form of it used in the Small Sample Study. Twenty-five cues were scored from the Self-Interpretation.\*

*Picture Reaction Test (PRT).* This procedure was somewhat shortened after the analysis of the Small Sample data. Since no published account of this test exists, it will be described here in more detail than the other procedures. The test materials are twenty-three pictures mounted on stiff cardboard cut to a uniform size of 10 x 12 inches. The pictures were cut from books and periodicals and include illustrations to magazine stories, photographs of varying artistic merit, and reproductions of paintings.

The subject is told only that he is to be given a test; he is then handed the first card. At the end of five seconds, the examiner asks him to hand it back (and takes it from his hand if the subject does not understand at once). The examiner then asks: "What was that a picture of?" After writing down the subject's answer, he says: "That's fine. I'll show you each of these pictures for just five seconds, and then I want you to tell me what it's a picture of, as you did here." After seventeen of the pictures have been described, the subject is told: "Now I'd like you to think back to the time when I handed you the first picture. See if you can remember everything that passed through your mind during those five seconds that you were looking at it." If after a short time he cannot recall the subject of the picture, he is reminded that it was a baby. If he mentions no affective reactions to the picture, he is

The method given here has been further developed and its general applications in psychodiagnostic work described (Luborsky, 1953). The original idea of self-interpretation of the TAT was derived from Bettelheim (1947).

asked: "Any feelings?" This further question is put to him if he does not spontaneously mention the topic: "Did anything pass through your mind in relation to the fact that it was the beginning of a test?"

After all introspections have been recorded, the examiner says: "Now I am going to ask you to look at the pictures again, and I'm going to add a few more to them. [Six more pictures are added.] Please go through the stack and sort them into three piles —those you *like*, those you *dislike*, and those you are *indifferent* to." If there is a question about the meaning or basis of liking, the subject is told that it is up to him. The sorting completed, the subject is asked to go through the pictures liked, indicating briefly why he liked each one. When he has done so, he is instructed to select "from four to six that you liked best of all." Finally, the subject goes through his dislikes in the same way, giving his reasons for disliking them and selecting the ones he dislikes most.

For the first picture, there are several kinds of data: original description; introspective recall; and, if the picture is either liked or disliked, reasons for the subject's affective reaction. Description and affective reactions are available for sixteen other pictures, but only reasons for liking or disliking in the case of the six added pictures. Total time of administration averages forty-five minutes. No less than 280 cues (not all scored on each case, of course) were listed in the manual for the PRT.

*Credentials.* These consisted of the following: all correspondence between the subject and officials of the School prior to his arrival in Topeka for assessment; letters of recommendation (generally three) he had been requested to supply, which sometimes included information about how he was doing as an intern; application forms (U.S. Civil Service Form 10-2852, Application for VA Residency, and Application for a Fellowship in the Menninger School of Psychiatry) and, in many cases, a brief statement by the applicant telling why he was applying for psychiatric training and something of his professional plans. Since there was a number of gaps in the relatively objective background data supplied by these forms, we devised a Supplementary Face Sheet to cover them, and one was filled out by each applicant at some time during his three days in Topeka.

#### PREDICTIVE RATINGS THAT WERE MADE

As was already mentioned, predictions about psychiatric work were made on four levels of generality-specificity. Most general was the rating of *Over-all Competence*. Next came the four general aspects of the resident's work which, with Over-all, we designated as major variables: *Competence in Psychotherapy*, *Diagnosis*, *Management*, and *Administration*. Still more concrete were fourteen of the

twenty-four "specific aspects of work" already referred to; finally, sixty-eight statements were taken directly from supervisors' comments without any generalization. (The last two types were grouped together in a form that is reproduced in Appendix 10.9.) The fourteen minor variables were:

- A1. Ability to Show Warmth.
- A2. Ability to Inspire Confidence.
- B. Empathy.
- C. Freedom from Overidentification.
- D. Interest in Patients and in Psychiatry.
- E. Spontaneity vs. Inhibition.
- F. Acceptance of Responsibility.
- G. Tolerance for Patients' Aggression.
- H. Capacity to Tolerate Stress.
- I. Freedom from Hostility toward Patients.
- J. Firmness, Authority vs. Vacillation, Domination, or Overpassivity.
- K. Sensitivity.
- L. Judiciousness vs. Impulsiveness.
- M. Ability to Communicate with Patients Effectively.

#### THE STAGES OF ANALYSIS FOR EACH JUDGE

The basic design of the Predictive Study is summarized in Table I-10.1.\* At the first stage of analysis, each judge began a case

TABLE I-10.1. OUTLINE OF DESIGN FOR ANALYSIS OF DATA IN THE PREDICTIVE STUDY

STAGE OF ANALYSIS	JUDGE I (Initial Impressions)	JUDGE II (Initial Impressions)	JUDGE III (Initial Impressions)	JUDGE IV
Stage 1	PRT	TAT	Interview	Rorschach
Stage 2	Rorschach	Self-Interp.	TAT	Credentials
Stage 3	TAT	Interview	Self-Interp.	PRT
Stage 4	Self-Interp.			
Final stage	All data <sup>a</sup>	All data <sup>a</sup>		

<sup>a</sup> Note: At this stage, all procedures listed in the table were available to the judge, plus the Wechsler-Bellevue and Word Association tests, and the Supplementary Face Sheet.

with only a single source of data available to him. For Judge I, this was the Picture Reaction Test; for Judge II, the Thematic Apperception Test; for Judge III, the recorded and transcribed Interview;

\* This design is derived from the one used by Horn (1943). See Appendix 10.8 for further discussion of Horn's work.

and for Judge IV, the Rorschach. After reading the protocol (and listening to as much of it as he wanted, in the case of the Interview), the judge then set about scoring the manual on special scoring sheets that were provided for the purpose. He did not have to make any notation about his mental processes in this scoring unless there was something unusual about it. (E.g., one judge occasionally noted that he was scoring a certain content cue from the TAT because it fitted the description given in the manual, even though he felt that it did not have the positive significance it was supposed to, and gave his reason for thinking so.) Such notes he recorded on a blank pad; they were filed along with the scoring sheets and protocols which he had marked up.

As he scored a cue, the judge also posted it on a second scoring sheet, on which ratings of personality variables were made. In the manual, each cue was accompanied by an indication of the personality variable, or variables, that it presumably indicated. Thus, if the subject said of the boy in card 1 of the TAT that he enjoyed playing the violin, the judge marked a plus on that TAT Content Scoring Sheet and also posted a plus on the personality variable sheet on the line by Internalized Interest in Work, which according to the manual was indicated by this cue.

When he had finished scoring the manual (or, in the case of the TAT, the two manuals—separate ones for Content and for Formal Aspects), the judge, therefore, had one predictor, the cue-sum, and a summary sheet peppered with pluses, minuses, and half-scores. In addition to those dictated by the manuals, he had also posted pluses and minuses for all other features of the test that he thought amounted to positive and negative indications of the personality variables. (These were placed in parentheses to distinguish them from the others.) For each such judgment he had to make a note on his pad of its source and explanation. He could also mark up the protocol to indicate its significant portions and his interpretations of them; and actually the making of notes became streamlined to little more than this for many frequently-made interpretations.

The judge then computed the algebraic sum of pluses and minuses for each of the thirty-two personality variables and rated it on a scale from 1.0 (undesirable) to 6.0 (desirable). At the same time, he explained and justified these ratings on his pad. Note that he was rating the personality of the applicant as of the time he came to Topeka to apply, though the predictive ratings about to be described were aimed forward in time.

Proceeding next to predictive ratings, the judge considered the minor variables, rating each on the same 51-point scale that was used for all predictions; 2.1 was supposed to be the lowest rating given to an acceptable applicant, whereas 2.0 and lower indicated rejection. By each rating he made notes of the personality variables he considered particularly important in influencing his judgment on it and recorded any other reasoning involved. At about the same time (he could do the most specific statements before or after the minor variables), he considered the group of statements under each minor variable and marked whether or not, in his judgment, each one applied to the subject. Every such rating was an attempt to predict a final criterion rating, which was to be based on the resident's final two years.

Up to this point, the judge was free to omit the rating of any personality variable or specific aspect of work if he felt that he had insufficient basis to form a judgment about it. But when he came to the final set, the major variables plus Liking, he was required to rate each at all stages of analysis. Again, every rating had to be accompanied by notes explaining it.

Somewhere along the line, between scoring the manual and starting the next stage of analysis, the judge wrote his sketch; its timing was up to him. It was an interpretative summary, often including diagnostic and genetic hypotheses, but focused on the attempt to visualize the subject as a personality involved in the work of a psychiatric resident. The sketches, therefore, often included qualitative predictive statements on matters not covered by the ratings.

All of the above brings us only to the end of Stage 1. The judges now proceeded to the next stage of analysis. Judge II, for example, turned to Self-Interpretation of four TAT stories which each subject had written. He proceeded just as he had done in Stage 1, making the same ratings, etc., with one difference: He could now base his judgments on both the TAT and the Self-Interpretation, since the stages were cumulative for all judges and each completed all stages on one subject before turning to the next.

With the repetition of the same procedure for Stage 3, Judges III and IV were through. The original plan was that all judges would go through the entire file on all subjects, but some did not have the time.\*

\* Judge III came to Topeka for the summer of 1950 to do his series of cases. After we had kept time records on the first few cases, we saw that he would be unable to get through if he did a complete analysis of all of them. Conse-

Our original intention was to use enough judges so that there would be two for each potentially important battery of tests, but our funds were insufficient to carry out a larger design. Consequently, we used the compromise design reported here. It has the flaw that there is so much confounding of judge and battery that we cannot separate the two. Each judge could be compared with one other at one or more stages, however; for example, Judges II and III at Stage 3. Moreover, each manual was scored by at least two judges.

We had initially planned for all judges to start each subject's analysis by reading the Credentials file and making a series of judgments based on Credentials alone. It was considered desirable because we knew that data of this kind were likely always to be on hand and to form the basis of the first screening before any elaborate assessment took place; therefore, in order to determine how worthwhile any more expensive technique or battery might be, it would be good to be able to see how much it contributed in addition to, or in correction of, what was gotten from the Credentials.

There were the following arguments against this proposed procedure: The Credentials tended to be one of the worst sources of contamination (see Appendix 10.9) and gave us very little that we could base any legitimate conclusions on; and they added a series of ratings which took up a good deal more valuable time than they seemed to be worth. Neither of these arguments ought to have affected an ideally designed study, but under the circumstances they were cogent enough so that Credentials were analyzed first on only a small series of cases.\*

---

quently, it was decided that he should do only three stages and make a minimum of explanatory notes. Judge IV was really two people, both graduate students of clinical psychology with a couple of years of clinical and research experience. Each did about half of the cases. Since a correlational check showed no appreciable difference in validities, we have treated them as a single judge throughout.

Starting with the cases of Class VI, Judge I began each analysis by going over the Credentials. This was done for two reasons: first, in order to make his ratings at Stage 2 more comparable with those of Judge IV; second, in order to see what could be obtained from just the Credentials, which in the Michigan project had turned out to be such an unexpectedly good predictor of competence as a clinical psychologist. This constitutes a minor deviation from the design as indicated.

## A SPECIMEN ANALYSIS OF PREDICTIONS ON ONE SUBJECT

THE COMPLEXITY OF OUR SECOND DESIGN MAKES IT DIFFICULT TO ABSORB its details in a generalized presentation. Therefore, we shall use a concrete case to illustrate our procedure of making predictive analyses.

Let us introduce Dr. Abbott in the same way that we first saw him ourselves. In May 1949 an applicant to the Menninger School of Psychiatry was scheduled for several hours of tests and interviews with members of the Selection Project staff (though he did not know until the last of these was over that they were not a part of the regular set of procedures in terms of which he would be accepted or rejected). When he showed up for his appointment with Judge II, he was given the TAT and Self-Interpretation. Right afterward, the psychologist noted his first impressions and incidental observations. The notes describe the applicant briefly: "Dark hair. Thin but wiry" and go on to include Dr. Abbott's remark that he had seen Dr. Yarrell, of the staff, to whom he had talked about his desire to get into psychotherapy immediately. "Dr. Yarrell told me that I shouldn't be in such a hurry to get started—I think he is right." The impressions record his answers to questions about where he had gone to medical school and his internship: "In the Army; I'll be out in September and ready to start here in October." He commented that he had heard of the TAT. Judge II concluded his impression: "Serious per-

sonal problems; now I think he would be a 2.3 [on a scale of 1.0 low to 6.0 high, with 3.5 as average]; after therapy he would make a 4 at least, or more."

Judge I wrote a similar physical description but continued in a more evaluative vein: "Somewhat breezy manner. Never felt wholly convinced by him—doesn't come through as a warm and sincere person, possibly because of anxiety, which he avoids. Probably too aggressive. Much on both the side of assets and of liabilities." Judge I's impressionistic predictive rating of Over-all Competence (using the same 6-point scale) was 4.

#### THE PREDICTIVE PROCESS

A little over a year later, after he had seen a good many other applicants and had completed the revisions of his manual, Judge II was handed a folder marked "Predictive Study Case 16"; it contained only a typed TAT, with black masking tape over the subject's name and the date he had been tested. The judge knew that the subject could be any one of about forty-five acceptees; as he looked at the stories, he had no sense of recognition, no recall of an associated face or conversation. He had given the TAT to about a hundred applicants during the preceding eighteen months and did not even know which of them had been accepted by the School.

#### STAGE 1: THE TAT

Here are the stories, the first four of them having been taken down verbatim as Dr. Abbott spoke them aloud, whereas the remainder were written out by the subject. In what follows, the bracketed interpolations are Judge II's marginal notes made as he went over the test; the italics are aspects that struck him as significant.

Can I take as long as I want before . . . ? [Concern about following directions. Desire to have as good a chance to do as well in the situation as possible.]

Story 1. (Smile.) Well, this is a young lad who has been started in violin lessons by his folks, who has been put in a room to practice (smile; laugh), who was not happy about it at all. In fact *he has almost fallen asleep* in the attempt to get away from the necessity of playing. In fact he would rather be outside. The mother and father are sensitive people who both enjoy good music who have yet *antagonized* the boy in such a way that he has not accepted the idea. [The

smile is to reassure about the antagonism.] He's been told many times that even though he dislikes it now he will like it later, but he's not convinced. Although he certainly will be; I like the boy (smile). [Inappropriate emotional reaction; poor Emotional Appropriateness.] My inclinations are strongly to feel sympathetic and to project myself into the situation. [Self-reference.] He uh, might well be waiting for some of his pals to rescue him, uh, since he has not yet learned or feels afraid, of taking positive steps toward getting himself out of the situation. It seems that he's thinking a great deal (laugh) how he'd rather be outside playing ball, cops and robbers or something like that. The parents are discouraged; due to their lack of understanding of their boy, they are at a complete loss as to what to do to change his ideas. [Negative, stubborn.] You say you want this projected into the future? (Yes.) Well, as he grows older, he, uh, learns much as he has been told that he begins to enjoy it; he will, if he is able to become technically proficient, be an excellent violinist but certainly not for anything more than his own pleasure (shrug). That's enough. [*I.e.*, he'll repay them by not giving them any pleasure.]

Story 3BM. . . . . Well, I get several ideas. [Obsessive.] It appears to be a young boy who, for some reason or other, is in a waiting room, perhaps in a doctor's office. Well, might even be a bus station; he's uh, *become very tired*, there was some emotional upset that's occurred. He's *fallen asleep* in this very awkward position. [*Cf.* first underlining, story 1.] The uh mother, has brought, him here, well, let's say to see the doctor and he's objecting strenuously, and a scene followed, and he threw a *temper tantrum* and here he is. (Pause. Laughs.) Golly! . . . [Affect expression; poor Emotional Control.] The small object near the bench was at first impression a knife in a pool of blood, but it's easy to change that impression to a toy gun that the boy had with him. . . . Part of the, the boy's, uh, golly! . . . problem is that he's been brought up in a family that very strongly desired *a daughter*; I get the feeling particularly from his large hips, that he may feel tremendous *doubts about his masculinity*. [Poor Heterosexual Adjustment.] Well, I said the most important part. (Stops.) That's — (What's the plot?) *I don't get a plot; all I get is an emotional reaction*. [Poor Emotional Control.] (These are all supposed to be stories, with a beginning and an end.) Well, due to these problems, he has been taken to see a physician who is supposed to help the boy with his problems, the boy wanted to uh, feel dependent, and yet he objected violently which led to the tantrum; while the parents are still talking to the physician he ended up in the waiting room. The outcome will probably be that he will grow up and *through a lack of proper sort of help* [poor Feelings of Security] will continue acting in this irrational and unsatisfactory way in solving his problems. Now you got a story. [Poor Handling of Hostility? Now you made me comply.] You might

add a note that again I identify with this individual in the way I've already expressed. (Not necessary.) I realize it isn't necessary. It's perfectly obvious. [Good Psychological Mindedness, though questionable.]

Story 4. . . . The present situation is that this married couple have reached uh, a, point of separation, nearly, because the husband has found, that he is in love with another woman and feels so strongly guilty about it, that he is, uh, his wife has learned of it through his unconsciously determined *clumsiness*. [His impulses get out.] She, on the other hand, is very reluctant to let him go, because of her own feelings of love for him. It's amazing but *she's not jealous*. [Defends through denial.] The outcome will probably be that the husband, feeling a *strong dependency* on her [somewhat inappropriate psychologizing] as an image of his mother, will turn back, uh . . . and re-accept the security of the family situation. But perhaps *trying in vain to forget* [attempts to use repressive defense] his past escapade. One gets a strong feeling from the *sturdiness of his features*, [attempt to infer from the picture] that he may be able to resolve his own problem and thereby establish a happy marriage, a satisfactory and mature marriage particularly since his wife has contributed so much to its establishment already. Have we got a beginning, middle and end? (Yes.)

Story 15. (Long delay.) It's a very old and sad *colored gentleman* [Ethnocentric Prejudice?], who after a long life of disappointment and unhappiness, has uh, lost his wife, and feels as though his own life were about to end. He has come to the graveyard, to her grave, with very uh, strong wishes that he could join her. He will probably do so in the very near future; there were no children in the family, which has only increased his *loneliness*. I don't know, do we have a beginning, middle and end? I guess so. [Doubt; indirect question.]

Story 6BM.—The present situation is that the woman's husband, and boy's father, has recently died. He had been a fine *old gentleman* [cf. story 15, first italics; southerner!] and he and his wife had been happy together, having raised a family of several sons and daughters. The boy, who was away has returned for the funeral and is now about to leave his mother *alone*. [Cf. loneliness in story 15.] She feels, of course, quite *lonely* and yet not completely lost. The son, knowing his mother's grief wonders what he can do to help her through the weeks ahead. *His own inadequacy bothers him* [poor Self-Confidence] although he has provided for her financial security and tried to comfort her. [Good Nurturance?] In some ways he feels guilty about leaving her, yet he knows full well that he must and that *time will heal all the wounds*. [Trite slogans.] The mother expects nothing more from her son than he has already given and though she only partially understands his problem is yet dimly aware and that will help

her [Nurturance?] solve her own problem. [Defense: guilt leading to resolution.]

Story 7BM. The present situation is one in which a *hostile, aggressive son*, who has been remarkably errant has returned home to his wise and kindly father and who resents that kindness because of *his own inability to return it*. [Poor Nurturance? Cf. story 1.] The life between them has been one fraught with disagreements and attempts on the part of the boy to get away from his father. Finally when he left home he found himself a rather *total failure* [poor Self-Confidence] and his own *dependent needs* have driven him back. The father, who understands the boy's reaction because it has been his own many years before, will do his best to give the boy the understanding and feelings of being wanted that he will need to meet his own problems. I expect that years will change the young man and that with age he too will be faced with similar situations and then *play the role* that his father is now playing toward him. [Façade? Good Self-Confidence?]

Story 12MF. *It is possible that* [circumlocution] this is a young married couple and that the, the, the, young man has recently discovered that his wife is a prostitute. She has delighted in flaunting this idea at him and is doing so now. His own reaction is one of shame and *inability to face the situation*. His impulse is to run away but he is for the moment unable to do so. His *own past has conditioned him against* the sort of woman his wife turned out to be and he feels *bewildered*. [Psychologizing.] The wife, with *he* [?] face turned against the wall, is not completely playing her *role* [*i.e.*, stereotyped role; Façade?] either for she feels a few pangs of remorse at what she has done. It is probable that the relationship will break up with the young man leaving her and getting a divorce and starting out afresh. [Improbable kind of situation. Implicit that one can't know women well—they're liable to turn out to be evil creatures (prostitutes). Then it's a black and white situation—just calmly sever the relationship and "start afresh."]

Story 8BM. This is a situation in which a young boy, out hunting with an older man who has been a good friend of his, has accidentally shot the man with the rifle pictured at the left. *The full awareness of what he has done has not yet reached him and that, as well as the knowledge that it was an accident* [defenses], keeps him from feeling great remorse. The surgeons are operating on the man to try to save his life, and will probably do so since the rifle was of small caliber . . . and the boy who has enjoyed their past experiences together will be able to go out with him again and hunt feeling unafraid. The boy's own life has been happy and he has been able to withstand the shock well. The man who was wounded understands the situation as an accident only and *will have no fears either*. [Defense.]

STORY 11. (There was a printed instruction to tell a fairy story fastened to this picture.)—Once upon a time, way at the bottom

of a deep sea, lived three old witches [mother figures?] and although that was an odd place for them to live *they did not mind*. [Probably an effort at whimsy; if not, it has an odd flavor.] They lived in a great castle which had been built long before and they stayed to themselves most of the time. Occasionally however they wandered out to the floor of the ocean where they encountered great fishes and *crocodiles* and they wandered around the vast floor of the ocean. One day however one of the witches got restless and decided to go to the world on top of the ocean and the other two witches objected quite strongly—so outside the castle there was a big fight with the two witches trying to keep the third witch subdued and at the bottom of the ocean. There was, however, a long narrow rock ledge which led from the castle to the top of the ocean and dry land and the third witch, after fighting furiously, escaped and started on her way. She fought all sorts of fights along the way with various monsters that inhabit the ocean but soon the water began to get light and she ran pell mell toward the surface and soon beheld the light of day. Whereupon she ceased to be a witch any more and became a beautiful woman. [*sic*] [The need to escape and finding yourself also unable to act is a common dream experience.]

Story 16. (Another printed instruction told the subject to imagine a picture and then tell a story about it.)—This picture is a wooded area beside a river flowing peacefully along. The trees are large and green with foliage. In a small clearing at the center stands a young man, handsome and strong, with two companions who are wearing less *pretentious* clothes than himself and who are *jesters*. The young man is soon approached by a messenger saying that the king's daughter has just been captured by a *crocodile* [*cf. story 11*] man and taken to his magic cave at the *bottom of the river* [*cf. story 11*] . . . and that the king is offering a reward of money and her hand in marriage to any young man who can rescue her.

The young man in the picture immediately accepts the challenge and after girding himself with various *devices for killing* the crocodile man he sets off for the bottom of the river to find the magic cave. After a long struggle and a great battle with the crocodile man he is able to rescue the beautiful princess and carry her back to her father. After that he marries the princess and they raise a large family and live happily ever after. [Unusual to make this a fairy story—makes it into a play. He brings in a lot of things which are of no use in the story, e.g., jesters. Arbitrary?]

Story 12M. This is a situation in which the young boy has been hypnotized by the doctor who has had him lie on the couch and has induced him to go to sleep. The reason for this is because the young boy has had a great deal of difficulty at home and has *become a problem to his parents and himself*. This is one of methods they are using to cope with the situation. Actually the boy *seems quite peaceful* which cannot be attributed solely to the fact that he is hypno-

tized. [*I.e.*, his improvement is genuine, more than suggestion.] It seems entirely possible that he will be able to express some of his feelings for his parents to the hypnotist and feel much better afterward, when he awakens which may be soon if *the gesture of the hypnotist means anything*. *The doctor is quite satisfied* with the results and pleased that he has been and will be able moreso to help the lad. [But is the lad?]

These eleven stories were the raw materials for Judge II at Stage 1 of his analysis of the case. As he read them over, he underlined things that caught his eye and made marginal notations. But the first main task was to score the two TAT manuals. (Figure 11.1 \* presents the Scoring Check Sheet for TAT Content; Figure 11.2 the one for Formal Aspects of the TAT; the pluses and minuses are the scores given by Judge II.)

Once the cues had been scored, the next step was to post the various pluses and minuses on the sheet listing the thirty-two variables of personality, which the manual cues were intended to measure. Figure 11.3 presents Judge II's Summary of Indications of Variables in the TAT. After casting up the sum of positive and negative cues relevant to each variable, the judge made a "freehand" clinical rating of the strength of as many variables as he thought he could rate.

Here are the comments and explanations that Judge II recorded for the variables he rated:

3. Need to Help: He is concerned about his efforts at nurturant behavior; seems anxious about how nurturant those efforts are. In three stories (Stories 1, 6, 7) he shows blocking of giving to parents, although he manages (story 6) ". . . to try to comfort her" and adds that she really doesn't expect much of him.
4. Psychological-Mindedness: Gives some fairly good motivational description of characters but in general doesn't seem very psychologically-minded. Indulges in very average psychologizing.
8. Self-Confidence: Made difficult to evaluate by the presence of contradictory evidences. His identification-figures are often very dependent—have to rely on (or blame) parent figures or wife: Story 4; "sturdy features, establishes satisfactory marriage since his wife has contributed so much to its establishment already." On the other hand, his story outcomes *if genuine* (see notes re Genuineness) represent ability to find resolution, some adequacy to overcome the opposite feeling.
9. Self-Objectivity: Especially his self-references—not clear inference

\* These and all other figures referred to in this chapter are contained in Vol. II, Appendix 11.1.

to poor objectivity. Poor sensitivity to demands of the testing situation.

11. Capacity for Personality Growth: Especially feeling of self-defeat—although in most other stories OK. Inferences from pictures and questions to examiner about directions may indicate some rigidity.
12. Objectivity toward Authority Figures: Considerable conflict with them. Blaming them for difficulties; see TAT summary. Feeling of guilt for aggressive impulses and worry that they might be expressed. Strong aggression toward father in stories 7, 16—the crocodile man.
14. Empathy: *cf.* psychological-mindedness; some attention is given to mother and father's feelings though not especially insightful.
17. Mature Heterosexual Adjustment: Sounds immature, slightly over-dependent. Acting out of the same impulses as toward his mother.
19. Feelings of Security: Despite his unwillingness to give, parents seem to want to help out.
20. Adequate Emotional Control: Some undercontrol especially apparent at the beginning of the test. See summary.
21. Adequate Handling of Hostility. Fears that hostility will be expressed and reassures himself that it won't. See summary.
22. Emotional Appropriateness: Stories sound slightly inappropriate. Certainly his having to inform me what is self-projective is inappropriate. Also odd that he tells fairy story to story 16, as if he carried over the previous mental set. Also his great use of denial, e.g., "she wasn't jealous."
24. Freedom from Stereotypy of Thought: Although original, it is not an achievement, largely deviant.
25. Perceptual Sensitivity: Tends to distortion, especially under the impact of strong emotional response.
29. Genuineness *vs.* Façade: Emphasis on people's playing roles, as if they don't act their real selves but rather some stereotyped role—e.g., the colored gentleman, the prostitute, and things happen according to slogans ("time will heal all wounds"). (Perhaps connected with simple hysterical defense structure? Although defense is also through withdrawal from the situation.)

The "summary" referred to began as a collection of evidence bearing on the issue of emotional control:

Lots of smiling, laughing. He is under tension. It also is a way of revealing that he perceives the self-reference in his TAT stories. Stories are poorly organized, lots of inappropriate interjections of his personal reactions to the task and questions about procedure.

It then continues with a scrutiny of the reactions to stress in each story (omitted here since they are merely brief restatements of the story events), summarized as follows:

*Reacts to stress by:* Inhibition of thinking, repression, withdrawal: story 1, almost falls asleep; story 3, becomes tired; story 4, attempts to forget; story 13, inability to face the situation; story 8, awareness not yet reached him.—

Blaming much on parents; they usually set up the stress situation: story 1, put in room to practice by his parents, who antagonize the boy; story 3, brought to doctor by parents; story 12, difficulty at home, became a problem to his parents and himself. Blocked nurturance towards parents: story 6, can't give to mother; story 7, hostile aggressive son, resents kindness because of inability to return it.— Or by blaming self: story 4, guilty, clumsiness.—

Is able to reassure self about the bursting out of his impulses by believing that (a) someone else does not feel impact of them: story 6, young man feels guilty about leaving; mother expects nothing from her son that he has not already given; story 7, inability to return nurturance to father, who understands; (b) and that he had no awareness, conscious intent, or responsibility: story 8, the man who was wounded understands it was an accident and will have no fears either; story 4, he guilty, she reluctant to let him go, not jealous (she learned of it through his unconsciously determined clumsiness).

Judge I had a rather different view of this applicant, but remember that he had already been through two stages of data-analysis (PRT and Rorschach) and came to the TAT with some well-crystallized notions about the kind of man he was dealing with. Here are his scattered summary notes on the stories:

In stories 1, 6, 11, and 16 there are references to friends, companions and families with several children—all of which is unusual enough to suggest that he comes from a large family, with some fairly close and important siblings, or that age-equals of another kind (friends) have played an important role in his life, probably a supportive one. At any rate, only in story 11 do they seem to be anything of a threat. (Good Social Adjustment with Co-workers.)

In story 1, he fears to be self-reliant (probably because it means rejection of parents?). The theme of dependence-independence is a common one too (stories 3, 4, 7, implicitly in 15 and 6) and is a conflict that troubles him. He is immature (*cf.* Rorschach) and in TAT seems surprisingly unconfident of self—probably related to this struggle for independence. And very much to his unconscious homosexual conflict—“doubts about his masculinity” in story 3.

As in PRT, he shows here that he can identify self with parents, see their point of view—but feels insecure in parental role, doesn't know what to do.

Story 4: Strong guilt leads hero to give himself away and *can't forget* this escapade: self-defeat? (Formal 9) despite final affirmation

that he's strong, etc. Strong need for security of mother-substitute's love: Feelings of Security, +? or —?

Story 6: Conscious feelings of inadequacy and lack of manliness, depressive tone—may be his main handicap. Like Dr. B? [a chronically depressed resident in the school] Cf. also story 7, complete failure. He emphasizes that people must solve their own problems—others can help, but not do it.

Story 13: This may be a rather projective fantasy. (Cf. PRT card 18.)

Story 11: The hero is *female*: very unusual, a witch (cf. Rorschach, mother-identification?) Changes to a woman—has he had fantasies of doing so? Cf. story 3. All this implies that likewise, a beautiful woman may really be a witch—or a whore (cf. story 13). You can't be too careful. Poor Heterosexual Adjustment.

Story 8: A good deal of denial: necessary to keep the aggression against father well underground—cf. authoritarian admiration on conscious level (poor Objectivity toward Authority Figures).

Story 16. Odd that he tells *another* fairy story—misunderstood instructions? Self-Interpretation will tell. Now hero is conventional handsome, strong male. Crocodile man may be father? as castrator or persecutor. Immature fantasy—he retreats to safe world of childhood irreality. Preoccupation with bottom of ocean and of river is *unusual*: suggests regressive fantasies (back to mother).

Parenthetically, it is interesting to note how much both judges were impressed with the *negative* features of these stories, failing to take seriously enough the implication (which now seems plain) that he was quite conscious of most of the conflicts they picked up and thus probably had more control over them than they suspected.

After putting down these impressions of Dr. Abbott, the judges had next to consider how the personality implied would function as a psychiatrist. (Again, a rating form was supplied, but we shall not reproduce the original form here.) Here are Judge II's predictive ratings, based on a blind analysis of the TAT alone (with Judge I's ratings, in parentheses, for comparison).

#### PREDICTIVE RATINGS

- A.1. Ability to Show Warmth: 3.0 (4.1).
- A.2. Ability to Inspire Confidence: 4.0? (3.4).
- B. Empathy: 3.5 (3.0).
- C. Freedom from Overidentification: 2.5 (2.3). Cf. Emotional Control rating; from the emotional display at the beginning of the test, with low Self-Objectivity in inappropriate self-references.
- D. Interest in Patients and Psychiatry: 3.0? (5.0).
- E. Spontaneity vs. Inhibition: 3.7 (3.3). See note by C.

- F. Acceptance of Responsibility: 3.0? (4.0). On the slender evidence of dependence on wife in story 4, difficulty in facing situation—bewildered excusing self from responsibility and guilt.
- G. Tolerance for Patients' Aggression: 2.3 (3.0). Cf. rating on Handling of Hostility.
- H. Capacity to Tolerate Stress: 2.5 (2.5). Cf. rating on Emotional Control.
- I. Freedom from Hostility toward Patients: 2.3 (4.5). Cf. Handling of Hostility.
- J. Firmness, Authority, vs. Vacillation, Domination or Overpassivity: 3.0 (2.8). Like F.
- K. Sensitivity: 3.5? 2.9 (4.5). His comments to me may be an example of tactlessness.
- L. Judiciousness vs. Impulsiveness: 3.0 (3.4). None of his stories sounds very impulsive. (The identification-figure does do impulsive things but then tends to withdraw or get bewildered and then, as an afterthought, put on an act, "play the role," of rational activity.)
- M. Ability to Communicate with Patients Effectively: 3.0 (4.0). Choice of words only fair; story organization only fair. Perhaps he will intellectualize in psychotherapy since stories contain psychologizing.

Since the general ratings are the principal predictors, and since it happens that Judge II did not record in very much detail how he arrived at his, Judge I's ratings and justifications are presented separately.

#### GENERAL RATINGS

##### JUDGE II

*Psychotherapeutic Competence:* 2.8. Apparently Self-Confidence and Feelings of Security are OK (although the opposites are strong, and I'm not certain which to believe). Possibly in superficial relationships his ability to act the role of the psychiatrist is of some use. Liabilities: easily feels victimized in situations and hence antagonistic.

*Diagnostic Competence:* 2.8. Poor Psychological-Mindedness; can throw a few psychological terms around not very appropriately.

*Management Competence:* 3.0. Cf. Psychotherapy.

*Administrative Competence:* 2.8. He is not an especially systematic person.

*Over-all Competence:* 2.8.

*Liking:* 2.6. Possibility of limited Cultural Wealth; inappropriate emotional display, poor Handling of Hostility; stereotypic typing of people.

**JUDGE I**

(Note: After studying the TAT, Judge I raised his ratings on all the following variables except the first.)

*Psychotherapeutic Competence:* 4.0. If he gets therapy for depression and ambivalent parent-attitudes, 5.0. He's better than I thought in ability to *understand* the patient (raised ratings on Psychological-Mindedness and Empathy) and can express himself pretty well (raised rating on Quality of Verbalization) but is more unsure of himself (lowered on Self-Confidence) than I thought—less able to conceive and carry out a line of treatment. More dependent on supervision.

*Diagnostic Competence:* 4.8. Raised on Psychological-Mindedness, Empathy and Quality of Verbalization—and his low Objectivity toward Dependent Figures isn't so much a matter of distortion (despite some Projection) as of *overidentification*.

*Management Competence:* 4.2. He's cautious enough, and *able* enough, to keep himself out of too great tangles despite his poor Objectivity toward Dependent Figures; also good Psychological-Mindedness and raised Self-Objectivity should help him to recognize when things are getting beyond him and to get help. He's practical in orientation, seems to like people, has better Social Adjustment with Co-workers, Feelings of Security and Genuineness than I thought, so despite his lowered rating on Self-Confidence, he should do a rather superior job of management, working closely with supervisors.

*Administrative Competence:* 3.6. His ambivalence toward father-figures is somewhat helped by a tendency to identify belatedly. Self-Confidence down, but Self-Objectivity, Feelings of Security, Stereotypy of Thought, Genuineness, and especially Social Adjustment with Co-workers are all up.

*Over-all Competence:* 4.2.

*Liking:* 4.0.

Judge II had now finished his first stage of analysis on one case. We shall not follow him in quite such detail through the remaining three stages but shall emphasize the changes in his evaluations of Dr. Abbott as he saw increasingly more of the assessment data.

**STAGE 2. THE SELF-INTERPRETATION**

When Judge II finished his first stage of analysis, the secretary brought him the following page of Dr. Abbott's interpretations of his stories to four specified pictures, written according to instructions. (The italics indicate Judge II's emphasis and bracketed interpolations correspond to his marginal notes.)

## SELF-INTERPRETATION

6BM. The entire story and situation as I interpreted it is a reflection of my own difficulties in understanding and adjusting to my own relationship with my own mother. The central conflict, having deftly eliminated any conflict with father by killing him off, is one of breaking away from maternal bonds and establishing a life for myself. The guilt (my guilt) extends from *hostility toward my mother* [cue I: negative description of mother] (which I have eliminated in the story by making the family happy as mine was not) [cue W: +] and it has been my frequent attempt to pacify this guilt by buying things for my mother and providing "for her financial security," in that way subtly insulting my father who has never been able to provide quite as much as mother wanted. The statements in the story about the mother at least partially understanding her son's problems reflects two things about myself—first that I wish my mother had understood my own problems, which she gave little evidence for, and secondly that I secretly hoped that my own mother was aware of the sexual attraction which I felt for her, an attraction for which I always felt guilty and which was unquestionably a form of competition with my father, whose manliness must have been a threat, though *I was unaware of this until about a year ago.* [cue V: development.] In the story, as previously stated, I have finally conquered the father by disposing of him, and that of course is a wish on my part. I make the young man in the story feel inadequate when this has happened which is a *reflection of my own doubts about masculinity.* [cue H: self, negative.] The meaning of the statement that time will heal all the wounds is obvious in the light of the above, indicating *still present dependent drives on my own part* [cue H: self negative?], dependency that can only be conquered by a gradual process of maturing. [cue H: self positive?]

7BM. I didn't manage to cover this up at all, though I confess that was one of desires while writing the story. [cue O: acknowledgment of change.] The reason is *simple*—that I have always felt a great liking for my father, that I have usually gotten along well with him [cue I: father, positive] and until adolescence conformed pretty much to his wishes. During adolescence however the family situation became *abusive* [word misused] and after a big row with father, who (significantly) was defending my mother whom I had *called a bitch, and I promptly left home.* [Sounds impulsive; poor Handling of Hostility.] The errancies committed by the young man in the story were my own, largely sexual in nature and mostly unsuccessful fulfillments. In my own life I ran away from these errancies by avoiding them (in the story the lad goes back to father). The father who in the story is understanding and helpful [father figure positive] was in my own case a succession of various helpful persons who have offered clues here and there which have led to a fuller understanding

of my own feelings and desires and wishes. [cue Va: development.] My optimism for the young man in the story reflects my own optimism regarding *my own anticipated solution to, among other problems, the resolution of this oedipal situation.* [cue T: recognition of problem; cue H: self, positive?]

13MF. Here again the oedipal conflict is quite *obvious*. The prostitute representing my mother and the shame and reaction of the young man resembling my own when I at last began to realize that I had such desires for my mother. Her unfaithfulness to me of course is represented by her sexual relations with my father. My own reaction was similar to that of the young man being "for the moment unable to do" anything. The statement that "his own past conditioned him against that sort of woman" reflects a situation I became involved in while in medical school in which I had unsatisfactory sexual relations with a girl who realized better than I at the time that she was playing the role of mother to me and in fact said plainly that she was "sick of being a mother" to me. The past conditioning is *obviously* the unresolved relationship with my own mother, prior to the relations with the girl. The ending of the story indicates (in the man seeking for a divorce) my own desires to resolve the situation [recognition of problem??] and be freed of the bonds of desire for my mother. [Somewhat too glib; cf. initial statements in this and the interpretation of 7BM that things are *simple or obvious*.]

16. I claim no originality in this picture or story, for it is a classical Siamese drama which I saw in Bangkok and which I enjoyed immensely. [Distinguish this from cue A, peripheral determinants.] Again the story has to do with the oedipal situation but has a different note than my own previous stories in that it is concerned with the resolution of the conflict as well as a statement of it. The young man is of course myself and the princess a representation of a young girl I would like to marry. The reward of money and marriage is the reward of maturity, of having successfully and masculinely and aggressively fought off the rivals (the crocodile man patently represents father and other suitors) and won the struggle [cue T: recognition of problem] which I myself *must win and expect to win*. [cue H: self-positive; cue T: determination.] The girl is the mother substitute and gaining her represents having thrown off the shackles of childhood, and my own dependent needs.

Even as he read over the Self-Interpretation, Judge II was making notes to help him in the scoring of the manual. (His complete scoring appears in Figure 11.4.) These scores bring out the fact that this is one of the best Self-Interpretations obtained from any applicant in our experience. It is not surprising, therefore, that Judge II

raised his ratings of fourteen personality variables on the summary sheet and lowered only one.

#### SUMMARY (CLINICAL NOTES ON SELF-INTERPRETATION)

After reading the whole Self-Interpretation I get fearful that I was much too harsh in judging his TAT. He has considerable sophistication in the interpretation of fantasy materials and considerable self-insight. I may have been wrong in minimizing the clue to Self-Objectivity I mentioned in the TAT—his laughter representing his own awareness of meaningfulness; he even asked whether I wanted to be told of the meanings. I minimized it because it sounded inappropriate. I also wonder about his mono-principle oedipal interpretation. It is slightly glib and too ever-present. I had the impression early in the work with the Self-Interpretation that too much of a mono-principle was a bad sign (*cf.* Dr. R [a Small Sample Low] with the theme of "dependency"—this man has some similarity to R) but I never could prove this was true because of too few cases. Also the amount of his insight is so great that either he has been analyzed or otherwise treated or has had considerable experience; it could have some pathological significance if he has not been in treatment. For the last possibility there isn't enough evidence: a few words are used slightly inappropriately and stiltedly, some stories are unusual in a deviant sense, in story 3 he is too quick to give an affect reaction and a fabulized overvalent interpretation of the gun as "knife in a pool of blood."

The writing is somewhat better in these than in his stories (Quality of Verbalization).

Perhaps considerable neglect of the inverse oedipal situation—love of father.

The predictive ratings show very clearly the influence of the Self-Interpretation—not only the positive signs seen in it, but the reinterpretation of the TAT in light of this additional material. Notice, also, that Judge II now thought he had enough information to make more specific predictions; the statements he checked as applying to this subject are listed under the relevant headings.

#### PREDICTIVE RATINGS

- A1. Ability to Show Warmth: 3.0-4.0.
- A2. Ability to Inspire Confidence: 4.5. *Cf.* Self-Confidence.
  - 2. Inspires confidence of patients.
- B. Empathy: 3.5.
  - 8. Shows ability to empathize with patients.

- C. Freedom from Overidentification: 2.5.
- D. Interest in Patients and in Psychiatry: 4.2-4.5. Cf. Internalized Interest in Work.
  - 18. Very interested in his work.
- E. Spontaneity vs. Inhibition: 3.7. Cf. Adequate Emotional Control.
  - 21. Able to be spontaneous with patients.
- F. Acceptance of Responsibility: 3.5-4.0. But on the other hand he left his home; he is determined to work out his problem. He can seek advice and get help from father-figures.
  - 25. Has genuine feeling of responsibility for his patients. (?)
- G. Tolerance for Patients' Aggression: 2.7-3.8.
- H. Capacity to Tolerate Stress: 2.6.
- I. Freedom from Hostility toward Patients: 2.5.
  - 41. Loses his temper with patients. (?)
- J. Firmness, Authority, vs. Vacillation, Domination or Overpassivity: 3.6. Directness of Self-Interpretation (an avoidance of being bewildered?).
  - 45. Can be quite positive and direct.
- K. Sensitivity: 3.5.
  - 54. Has considerable sensitivity to what goes on between himself and patient.
  - 56. Makes grave mistakes through tactlessness. (?)
- L. Judiciousness vs. Impulsiveness: 3.
- M. Ability to Communicate with Patients Effectively: 3.5. Cf. Quality of Verbalization.
  - 65. Can get his points across to patients. (?)
  - 66. Intellectualizes in therapy. (?)

#### GENERAL RATINGS

*Psychotherapeutic Competence:* 3.5. He can be direct, apparently can accept his aggression better than I thought. Good Psychological-Mindedness and Empathy, although the limitation of his Psychological-Mindedness is that it is likely always to be about oedipal-level material and too little else. Good Self-Confidence and Feelings of Security. His own main problem is his unregenerate oedipus complex, partly covering his homosexual problem.

*Diagnostic Competence:* 4.4. Raised ratings on Psychological-Mindedness and Empathy; sophistication (Quality of Verbalization); orderly, systematic, complete.

*Management Competence:* 3.8. See above remarks on Diagnosis; also, he probably manages the "role."

*Administrative Competence:* 3.5. Systematic, orderly, probably likes the role though with some anxiety.

*Over-all Competence:* 3.5. Liking: 3.5. My conception of him has changed much. [Here he mentioned a couple of residents in the Small Sample.]

STAGE 3. THE INTERVIEW

The third type of data for Judge II was the recorded Interview. It is both too long and too personal to reproduce here, and, in any event, the reader could not have access to all of what was available to the judge, since the records themselves were included and the judge could listen to as much of it as he wished to get further impressions from the sound of the voice. We therefore include here only the scoring of cues, ratings, predictions and notes on them.

(Figure 11.5 presents the scoring check sheet for the Interview, with Judge II's scoring and his few notes; see Appendix 11.1.)

It is plain from Figure 11.5 that the interview gave Judge II a somewhat negative impression of Dr. Abbott. This is confirmed by the fact that he lowered about ten of his personality ratings and did not definitely raise any of them. Only the ratings on which there was a change or new notes are listed here:

PREDICTIVE RATINGS

- A1. Ability to Show Warmth: 3.1.
- A2. Ability to Inspire Confidence: 4.0.
- D. Interest in Patients and in Psychiatry: 4.0. See notes on Objectivity toward Authority Figures.
- E. Spontaneity *vs.* Inhibition: 3.7. Although the undercontrol which shows as spontaneity isn't always an asset.
- F. Acceptance of Responsibility: 3.0. He specifically says that responsibility beyond his knowledge makes him anxious. [Statement No. 25: Has genuine feeling of responsibility for his patients, is no longer checked.]
- G. Tolerance for Patients' Aggression: 2.4.
- 32. Patients can browbeat him. (?)
- H. Capacity to Tolerate Stress: 2.0.
- 39. Patients' problems arouse his own anxiety [double-checked].
- I. Freedom from Hostility toward Patients: 2.4.
- 41. Loses his temper with patients.
- J. Firmness, Authority, *vs.* Vacillation, Domination or Overpassivity: 3.3.
  - 45. Can be quite positive and direct. (?)
  - 49. Reluctant at times to be firm when he should be.

GENERAL RATINGS

*Psychotherapeutic Competence:* 3.0. Kept up because of Psychological-Mindedness and Empathy and his own statement of behavioral initiative and having been able to help patients. But lowered be-

cause of poor Heterosexual Adjustment, Handling of Hostility and especially poor Emotional Control; self-defeat, continuing chronic maladjustment leading once to mild suicidal attempt.

*Diagnostic Competence:* 4.4.

*Management Competence:* 3.8.

*Administrative Competence:* 3.0. See F. He can be systematic but has antipathy to work when he feels pushed.

*Over-all Competence:* 3.2.

*Liking:* 3.5.

#### STAGE 4. ALL DATA

After completing his ratings on the third stage, Judge II received all other assessment data available on Dr. Abbott: Wechsler-Bellevue Intelligence Test, Word Association Test, Picture Reaction Test, the Rorschach Test, Credentials, Supplementary Face Sheet—even his own initial impressions. Names and other identifying data were concealed in all of this material.

Its bulk precludes our reproducing here this mass of data. There were no cue-scorings at this stage, either; it was devoted to a clinical integration of all of the material, and the only ratings made were on the personality variables and the “general ratings” on the major work variables. After making these ratings, the judge wrote a personality sketch in which he tried to sum up his major impressions of the applicant.

For purposes of comparison, we shall again include, at this stage, the ratings and the sketch made by Judge I, who also made an analysis based on the entire file of material. The comments on the personality ratings are those of Judge II.

#### PERSONALITY RATINGS

1. Freedom from Material Preoccupations: 3.3 (3.1).
2. Freedom from Status-Mindedness: 3.0 (3.1).
3. Need to Help: 2.8–2.2 (4.7). Difficult to elicit warmth from him.
4. Psychological-Mindedness: 4.2 (5.9).
5. Psychological Curiosity: 4.0 (4.8).
6. Internalized Interest in Work: 4.0. Perhaps too high—because of his fear of what may happen to him without the help he can get for himself from psychiatry.
7. Internalized Interest in Psychiatry: (5.3) [Note: The two judges each rated only one interest variable but happened to choose different ones.]
8. Self-Confidence: 3.5 (3.8).
9. Self-Objectivity: 3.4 (5.0).

10. Sense of Humor: 3.2 (4.4). I see relatively little ability at humor. His laughing is more a release of tension caused by some disturbing problem.
11. Capacity for Personality Growth: 3.4 (4.7).
12. Objectivity toward Authority Figures: 2.2 (3.7). Ready to get into fights with authority figures.
13. Objectivity toward Dependent Figures: 2.5 (2.5).
14. Empathy: 3.2 (4.4).
15. Freedom from Ethnocentric Prejudice: 3.3 (3.5).
16. Consciousness of Social Injustice: 3.5 (4.1).
17. Mature Heterosexual Adjustment: 2.0 (2.0). I have serious doubt about how easily he can overcome his severe sexual inhibition, especially his impotence.
18. Social Adjustment with Co-workers: 2.5 (4.3).
19. Feelings of Security: 3.9? (4.2).
20. Adequate Emotional Control: 2.0 (3.0).
21. Adequate Handling of Hostility: 2.4 (2.6).
22. Emotional Appropriateness: 2.0 (3.0). Much lowered because of my feeling that lots of his reactions have an unusual, odd quality.
23. Clarity of Thought: 3.0 (4.1).
24. Freedom from Stereotypy of Thought: 3.6 (2.7).
25. Perceptual Sensitivity 2.5 (4.2).
26. Cultural Wealth: 3.5 (4.6).
27. Quality of Verbalization: 3.5 (4.5).
28. Well *vs.* Poorly Internalized Ethical Standards: (3.4) [not rated by Judge II].
29. Genuineness *vs.* Façade: 2.9 (3.5).
31. Projection: (3.3) [not rated by Judge II. If there were no particular evidences for the three pathological tendencies listed, no rating was to be given].

#### GENERAL RATINGS

*Psychotherapeutic Competence:* 2.5 (4.8). Could be anywhere between 2.0 and 3.5.

*Diagnostic Competence:* 4.0 (5.2).

*Management Competence:* 3.5 (5.0).

*Administrative Competence:* 2.6 (4.0).

*Over-all Competence:* 2.7 (5.3).

*Liking:* 3.2 (5.0).

Assets: Psychologically-minded, conscientious.

Liabilities: Homosexual development of his sexual impulse; conflict with supervisors and people who are aggressive to him (he thinks); hostility toward women; overintellectualization and isolation of affects; readiness to feel victimized; impression he gives of immaturity, probably because of some inappropriateness of his behavior. Lack of warmth: He reacts to the demand to work with all sorts of feelings of victimization and threats to himself.

Schizoid Developments: Forgetting, withdrawal, lacks warmth, probably few friends.

Perhaps slight psychopathic tendency: role-playing, rejection of patients with psychopathic tendencies (in discussing patients he dislikes).

As Judge I went over the credentials and read the physical description, he made the following note: "Sounds like Abbott—Oop! It is! I wouldn't have rated him so high." The secretary had failed to delete the subject's name in one letter. Therefore, being contaminated, Judge I made no predictive ratings at the final stage, so his general ratings from the previous stage are given here. (This was considered enough contamination to eliminate Judge I's ratings, even though he had no direct criterion information and his casual observations of Dr. Abbott would not have contributed spurious predictive accuracy.)

#### PERSONALITY SKETCH: JUDGE II

The subject is a dark, slender man of twenty-five, with long sideburns, who comes from a small town in Utah in which he lived all the early years of his life. His parents are of first generation; his mother's parents were Dutch. His father earned a very small but steady income. His parents were devout Mormons and he accepted religious teaching until he was eighteen, when he decided that "it was totally preposterous." As a child he was always frail and had considerable tension and tremors. He had fears of the dark, of snakes and having to work. He had lots of fights with other kids and was usually beaten up badly. He got along well with older people but not aggressive kids. He was spanked often by father who used physical punishment on him even when he was in his third year of college for his drinking. Whippings were usually for not doing things and his reaction was passive resistance. His parents had a constant conflict between them. The mother was extremely unstable and would become extremely angry with the father. There were frequent scenes of physical conflict between them. The father was always considered by him to be more stable except when he was goaded when he would blow off steam. He admires his father's equanimity although he criticizes him for not understanding mother. The main values of the home were particularly set up by mother: do well in school, do not be aggressive no matter what the others do, behave properly in company, be religious, and avoid social vices. He believes that he is like mother, that he has some of her instability, some of what he calls her "reaction to authority" which is a fierce hostility to authority, and an ability to be analytic on problems that are not emotionally charged. He has always liked father and disliked

mother and says that now he doesn't dislike mother. Some of the violent scenes between his parents disturbed him greatly. On one occasion the mother brandished a knife at the father and the father knocked her down. During this time the mother called on the son to save her and the son felt extremely upset and unable to help. This theme comes through in his TAT stories. The parents were separated for a year when the patient was thirteen and were re-united later.

In my first impression on meeting him, I thought that he had serious personal problems that would prevent him from working at anything better than a barely passable level but that with help he might become fairly good. I still have the same feeling about him. I notice that his references mentioned this same two-sidedness in their evaluation of him. One of them mentions that he has good knowledge but that he is immature, another mentioned that he is hardworking, conscientious, but due to his youth he was outspoken and critical even on issues on which he had no factual information. The evaluator believes that this defect may be overcome with experience. The applicant has had considerably more previous experience than most and has a psychological sophistication that is far above average. He's also had some experience with the psychological tests.

His serious problems center about the conflict between his parents and his inability to handle the resulting impulses in himself. His TAT stories deal with the following problems: being compelled to do things by his parents; separation of parents; death of one of the parents; dependence, especially on the father, together with hostility to the father; need for escape from the mother and sister; getting help for his own problems which are a problem to his parents. In his Self-Interpretation his approach is analytic and psychologically sophisticated but represents variations on the single theme of oedipal conflict. Throughout his tests he shows considerable concern about his masculinity expressed in preoccupation with what characteristics are associated with men and what with women. His fearfulness and hostility to women are so great that he is undoubtedly going to have difficulty in his sexual adjustment unless something very drastic happens to him. He will probably repeat the same pattern as his parents, the sado-masochistic conflict, with him being masochistic most of the time.

In handling of his affects and impulses his ideal is to behave properly, to avoid becoming aggressive and to accept adverse situations with equanimity just as father did most of the time.

Occasionally or perhaps often, as long as he isn't emotionally upset, he manages to be fairly direct (as in the Self-Interpretation and the TAT) without a great deal of obsessiveness. This directness has as its alternative, becoming bewildered and confused, which he does on some parts of some tests. In the PRT he fairly frequently gets stumped by some discordant quality he sees in the picture. (E.g.,

in picture 6, "they look like Slovakian people who are having a great deal of fun, over what it is hard to say, it stumps me" or in picture 7, "it looks like a nightmare, seems peculiar that the colors should be so bright and gay.") This kind of perplexing dilemma is apparently the result of his need to have a consistent impression of what he sees, an exact statement of what is happening. The constant discordances that he sees are projections of the tremendous discordance in himself. (*E.g.*, in picture 12 he sees a nun who has wisdom and understanding "but it goes against my concept of these nuns.") *I.e.*, he projects wisdom and understanding, which is one side of his conception of older women and then can't reconcile the other side which is his hostility and inability to tolerate this hostility for them. This exactness, which avoids anxiety of not knowing what is going on, shows up in the Wechsler in being more exact than he needs to be, or in the Rorschach, in a great effort to be exact with sudden revising because of doubt. Part of the doubt is that he may not be responding adequately or completely enough; he makes such comments as, "I can't figure this out," "there ought to be something in here," and considers then all the possibilities.

In addition to avoiding anxiety through exactness or intellectualized explanations he very often reacts to stress by an inhibition of thinking, a kind of withdrawal from the conflict (especially clear in TAT stories 1, 3, 4, 13 and 8). In these he describes the identification figure as "almost falling asleep," "becoming tired," "attempting to forget," etc. In describing his mother's reaction to a struggle with his father, he says that she often retired to her room for several days and didn't take care of herself. He himself describes a "feeble suicide attempt" when his girl friend left him and adds that he can run away from conflict as well as anybody. Another way of self-protection is through extra-punitive blaming of his parents or blaming of authority. This has a somewhat paranoid quality as if he is being made the special victim of these people. Another way is through self-blame, as illustrated very strongly in the suicide attempt and more frequently in his readiness to be self-deprecating, to say when he does something wrong, *e.g.*, in the Wechsler, "that was stupid." In the interview, he reveals things about himself that might be self-defeating.

He uses several ways to reassure himself that his impulses won't come out and hurt anybody or hurt himself. *E.g.*, he reiterates in the TAT that someone else doesn't feel the effect of a certain impulse he has. Likewise he tries to reassure himself by saying that he has no awareness of an impulse, no conscious intent, therefore no responsibility for it. A more direct way of self-assurance is shown in the interview where he denies his anxiety on several occasions by saying toward the end that his anxiety was not as great as he anticipated, that he feels relaxed, or in describing his relationship to subordinates he says, "I never got upset about it" (*i.e.*, the enlisted men not doing

what they were told), or in describing his fearfulness as a child he says, "I don't think I shook."

There is considerable undercontrol expressed in all the tests (*e.g.*, the behavior at the beginning of the TAT, the constant exclamation of "Oh boy!"). This undercontrol is shown in the interview, when he says that he fears that his tendency to become anxious is one of his main liabilities. He has considerable anxiety in making decisions. He becomes upset by a child crying. This high emotional responsivity may be reflected in his Rorschach with seventy-two responses. However, if the Rorschach is taken as adequate evidence, he also manages to preserve control most of the time. The F+ % is high and form is used in most of his responses. Perhaps what shows elsewhere as intellectualization shows in the Rorschach by his difficulty in using a whole approach, *i.e.*, he concentrates on details as opposed to wholes, and even gives several *Do* tendencies. The Rorschach also implies that he has considerable sensitivity and some originality which is on a good form level. Another defensive mode might be called role-playing. In the TAT, several of his characters are described as filling a stereotyped role (*e.g.*, the colored gentleman and the prostitute). In the PRT he gives an example which shows that this role-playing is sometimes related to a denial of being afraid. He describes in picture 4 the husband's being horrified from the looks of his wife's new pajamas. Then he says, "He's obviously not really horrified, probably putting on a show for her benefit; the maid seems to be playing along with it."

I don't have a clear grasp of his motivation for psychiatry except for the very prominent one of need for help for himself. He feels tremendous haste about getting this quickly, lest something happen to himself. He felt considerably reassured about this haste by Dr. Yarrell who said that he shouldn't be in such a hurry, and his reaction was that Dr. Yarrell was probably right. He remarks, in his Supplementary Face Sheet, that Dr. Yarrell is the man he wants to be like as a psychiatrist. This suggests a second motive, a part of his ego-ideal, that of taking the role of a father who isn't disturbed by adversity. He feels that it is quite important for psychiatrists to have a masochistic quality. He also has a problem in the area of giving and getting love. His two great aims, he says, are to become famous and to be loved, especially the latter. [This report ends with a penciled additional note, "I probably have under-rated him."]

#### PERSONALITY SKETCH: JUDGE I

The subject is a dark intense man with a breezy manner covering a good deal of anxiety and tension. He is quite self-centered—not narcissistic in much of a self-adoring way; he does not glorify himself at all but simply looks at things a little too much from his own personal vantage-point. The result is that he seems to have little warmth despite a surface of pleasantness, and one feels uncertain

about his sincerity. What façade there is, however, is the relatively forgivable and unpretentious one of anxiety-avoidance; he does not posture, does not try to appear more important or mature or intelligent than he is—just happier than he is. Still there may be some strain involved in the maintenance of even this front, which he has at times tried to relieve by drinking beer; all of which doesn't make such a good surface to present patients in management contacts.

Yet under this apparently cool and self-centered front there is a good deal of real warmth and desire to help others. It seems very clearly to be based on an overidentification with those in need and a desire to act out the nurturance he wants to receive. But he is quite aware of this mechanism, consciously struggles with his dependent and passive cravings and wants very much to be loved. I think that he is likely to be able to control his overidentifying tendencies because of this insight, although a good deal may depend on his chances in establishing a good love-relationship while he is here. If he remains isolated from love, or if he gets involved in a repetition of the kinds of unhappy love affairs he has had in the past, he may have such a strong need for gratitude and affectional response from patients as to get into difficulties like countertransference and overidentification. I don't feel that his approach to patients will be so much distorted on the side of *understanding* them, since he appears to have a good capacity for empathy. Just how far he can go in appreciating feelings that he has not had himself in patients with very different kinds of problems I don't know, but it looks as if his projective inclinations are under fairly good control. His empathy does seem to be to a large extent of the involuntary, forced, vulnerable type; his own latent depression is easily stirred up by such moods in others, and probably his anxiety too.

One very definite asset is his excellent psychological-mindedness and curiosity—a strong desire to understand people and himself, and a real flair for such understanding. Somehow without a great deal of background he has been able to absorb and grasp the essentials of psychoanalytic thinking for himself and to apply it unusually well to the mastery (through understanding) of his own problems. His own masochism, which he points out is derived from identification with his father, probably helps him achieve objectivity of the kind that hurts; he may actually not do himself justice in his interview for this reason. He had to struggle as a child with a rather intolerable family situation: an extremely moody and unstable mother, who provoked fight after fight with her abnormally long-suffering husband and would sulk in her room for days afterward. To survive he had to develop what understanding of her he could. The trouble is that it was just too traumatic, and he couldn't live through such a childhood without serious scars. He needs treatment and has enough insight to recognize it; I hope he can get it here, though he may not. It is a tough character problem, with his maso-

chism going back to childhood—when he sought fight after fight only to be beaten again and again (apparently acting out the pattern of his family life). There are important areas in which his insight is blocked, but as a whole he has a remarkably good understanding of himself.

I don't see the pattern of his identifications and his handling of the Oedipus complex very clearly. It's clear that there are important elements of both parents in him, that he has too much identification with a hysterical mother along with some physiological slowness of sexual maturity that made it difficult for him to build up a satisfactory masculine ego identity. He seems to have had a good and loving relationship with his mother during the first few years of life, a basis for apparently rather good basic security and ego strength; his good natural endowment of intelligence must be counted there too. To some extent his feminine identification, his latent homosexuality, seems to be ego alien; he apparently has little insight into his projective mechanisms. It may be that such a stormy, sexually attractive woman as his mother was too threatening, and he had to renounce her for his father as his unconscious love object at first. His statement that he got along fine with his father and was very fond of him until adolescence is consistent with this; his projective tendency now to see father figures (e.g., Army) as violently unfair and all wrong, his unusual insistence in the Self-Interpretation on sexual feelings for his mother, are consistent with the interpretation that as adolescence approached and he did not develop as fast as his friends into adequate manliness, he doubted his masculinity all the more. His father, as love object, then became too threatening, and had to be made persecutor, while he had to exaggerate consciously his heterosexual tie to his mother, to deny the homosexual feelings.

At any rate, he was able to rebel successfully against his parents, first against their values and then openly defying them and leaving home. This took some guts, and along with some other indications convinces me that he does have enough basic self-confidence, even though he also has a lot of self-doubts. The latter are closely related to his conscious feelings of sexual inadequacy, which in turn he relates to an incestuous conflict. He may be right, but his failure to establish a real love relationship may also be due to an unconscious homosexual problem, and basically an aggressive problem.

There are a few signs of his wanting to turn away from the frustrating quest for love and take power and prestige instead. But they do not seem very important or ingrained. His desire for money and thinking in terms of it are naturally enough related to his early poverty, but again this striving doesn't seem as important as love-seeking. His need for love is so intense, and his masochism also so ingrained, that he did make a kind of suicidal attempt when he lost the closest thing to a love-partner that he had had. In doing so, he

tried to drown himself, an act that was heavily overdetermined. The hysterical acting-out quality, which he recognizes, was an expression of identification with his mother; the masochism, his father; the choice of water, a desire to return to the good mother, the satisfying but perhaps too engulfing relationship he had with her as a baby. Somehow, he has to go down into those waters again to be reborn, this time to shake off his mother's clutch, overcome his regressive and dependent longings and become a man; let's hope he does it by psychoanalytic treatment.

As indicated, the handling of hostility is one of his central problems. To too great an extent, he turns it against himself and strikes at others through self-defeating passive-aggressive mechanisms which he recognizes that his father used against his mother. But he's not meek—at least they never "broke his spirit"; he isn't a mealy-mouthed, façade-ridden person crawling with reaction formations. He's full of conscious conflict, irritable and at times hypercritical although never apparently really getting his hostility expressed. Right now he is more critical of his mother than his father; he needs to become more aware of his hatred for his father, probably on the score of his own lack of manliness. That is, he is perhaps unaware that his father was a castrator in some sense despite his apparent mildness, when his Golden Age ended, and later on deprived him of manliness by being insufficiently manly himself.

I think it will take a long time, a long analysis and a period of maturing and establishing a good marriage before he can become a really good psychiatrist, but I think that he does have the essential abilities and that his character neurosis is curable. His ability to develop constructively he has already demonstrated.

#### DR. ABBOTT AS A PSYCHIATRIST: CRITERION INFORMATION

With these ratings and qualitative impressions, our attempts to predict how Dr. Abbott would do as a psychiatric resident ended. We knew that he had been accepted, and nothing more; even this fact did not have a marked biasing effect on our ratings, for we knew that some of the accepted candidates would turn out badly—some might even leave the field of psychiatry. Indeed, we were hoping to be able to identify such cases who had not been screened out by existing procedures. Both Judge I and Judge II expected this candidate to succeed in the school. Judge II's final Over-all rating was somewhat below the average prediction he made for the entire group of sixty-four subjects; it ranks forty-seventh. Nevertheless, any rating above 2.0 was considered a prediction of success in residency. Judge I expected an eventual performance a good deal above the average: he

ranked Dr. Abbott fourth out of sixty-four on Over-all Competence.

The degree of success or failure in predicting about Dr. Abbott cannot be considered in any way typical, since one case is not a sample. Nevertheless, it will be useful to complete the story, exemplifying the kind of criterion data at our disposal and the way we used them.

Our principal criterion, Supervisors' Evaluations of Over-all Competence, puts Dr. Abbott near the top of the group: sixth out of sixty-four. His contemporaries did not think quite so highly of him; his rank on Peers' Evaluations of Over-all Competence falls between nineteen and twenty out of forty-six, or just above average. Judge II thus came somewhat closer to Peers' Evaluations than Judge I, whereas the latter predicted Supervisors' Evaluations rather closely.

The final criterion leaves out a good deal of the story, however. Dr. Abbott started out well below average, receiving the following supervisors' ratings:

End of first year (Over-all): 1.3, 6.1, 8.3	Mean: 5.2
Eighteen months (Over-all): 3.7, 8.3	Mean: 6.0
Two years (Over-all): 3.7, 6.0, 4.8, 5.7	Mean: 5.0
Thirty months (Psychotherapy): 9.5, 11.8	Mean: 10.6
Three years (Over-all): 10.8, 9.6, 11.1	Mean: 10.5
(Psychotherapy): 9.8, 10.8, 10.9, 10.9	Mean: 10.6

These ratings are on a scale from 0 (poor) to 13 (excellent). At the end of two and one-half years, he was not rated on Over-all Competence because he was doing only psychotherapy at the time.

The ratings show a picture of erratic and not very good performance at first, a slight dip at the end of the second year, and then a sudden spurt to a high and steady level of performance, with very little disagreement among criterion raters. What happened to bring about this dramatic improvement, which stamps Dr. Abbott as a "Late Developer"? For the answer, we must turn to the qualitative evaluations, which also tell why he was put into the subclass of Late Developers called "Neurotically Inhibited."

Dr. Abbott began his first year by working with chronically, and then with acutely, psychotic patients. He had not been at it long before he again felt the need for psychotherapy. In November of his first year, he applied at the Menninger Clinic for treatment. His Section Chief wrote about him at that time (in a routine evaluation for the School):

"An above average resident who has personal difficulties of which he is aware and who functions effectively in spite of his difficulties . . . is socially and personally aggressive but he is aware of this and handles it in an acceptable manner. Has tendency to resent authority, is impulsive in his judgments and makes mistakes which, however, he readily recognizes and doesn't become hostile about. . . ."

Early in spring, he entered treatment with a psychotherapist. At the end of this first year, the supervisor who gave Dr. Abbott his lowest rating had had no direct contact with his psychiatric work.

"I know him from seminars and groups," this supervisor said in his evaluation interview. "I think he is insensitive and rather cold, an aggressive personality with ethnic prejudices. I think he lacks sensitivity altogether, and I really don't think that this man should do psychotherapy. As a matter of fact I am rather doubtful that he should become a psychiatrist at all. Although some of his gross difficulties may be due to morbidity—probably they are—some of the smoldering anger and prejudice could probably be helped by therapy; but unless he changes I think that the aforementioned opinion remains valid."

The supervisor who rated him most highly also described his contact with Dr. Abbott's work as "quite limited—he has been seen on the acute service in conferences about every two weeks for two hours but only for a couple of months. Actual contacts have been no more than three or four. . . . So far as an opinion can be ventured, this fellow has a tendency to approach patients with his own preconceived notions and, to too great an extent perhaps, talk down to them, to be a good authoritarian with them and somewhat unfeeling of their reactions to him in interviews."

The supervisor who knew him best rated him right in the middle of the scale on Over-all Competence and not very different on any of the other work-functions. "I've known this resident for about four months and have had opportunity to supervise his work for about two months," he said. "He impressed me as being a rather impulsive, fickle and perplexed person. Tended to go off on tangents, get involved in situations needlessly. Nevertheless, his administrative management of patients was fairly good and his case reports were well written. He was inclined to be somewhat hostile, argumentative. This was reflected at case conferences and group controls."

At the same time, two supervisors reported on him as follows to the School: "This man has good understanding and is able to manage his ward and patients adequately; however, he is reluctant to follow established procedures and often has to be shown that there are good reasons for most of the things we consider good practice of psychiatry. He has a flair for the dramatic which may be turned

to good use later. He tends to overidentify himself with patients, but there has been increased awareness of this on his part and it is less of a problem than it used to be. He has the capacity to be a good therapist and is showing improvement." The other wrote: "In my seminar, he demonstrated that he is quite an avid reader, having read many things that were not specifically on the reading list. He tends to be dogmatic in his assertions and has some difficulty in broadening his perspective; however, he has made progress and in my opinion has made a reasonable adjustment as far as his psychiatric reading is concerned. He was always ready for a friendly argument and served as an excellent stimulant to the group to think. His performance in group was above average."

In summary, then, Dr. Abbott's intellectual assets began to show themselves rather quickly, but he got into various kinds of mild trouble through poor emotional control, difficulties in accepting authority, and defenses against anxiety, including an aggressive manner.

During his next six months, Dr. Abbott spent most of the time on the neurological service, fulfilling this part of his residency requirement. A couple of months of that time were spent on a small section for ill and infirm neuropsychiatric patients, where he had a chance to practice some psychiatry along with neurology. The consultant to that section described him as "somewhat argumentative, with more interest in theory than in practice. Yet he is able repeatedly to relate himself pretty well to patients."

His supervisor on this section said, in part: "He had a limited tolerance of aggression on the part of the patients. He was impatient with them, it was hard for him to take aggression. One or two of his patients were quite noisy and made even more demands. Patients like that caused him to become tense, and he had difficulty in keeping some kind of equanimity. I think he is a brilliant fellow—intellectually very keen, so that he really has no difficulty in understanding the theoretical aspects of psychiatry. I don't know, I thought that there was a tendency on his part to be somewhat depressed. I felt that he really didn't get intense pleasure out of psychiatric work. He likes to work with people who are ambulatory, who can respond to him. Not those whose sensibilities are so disturbed organically that they can give very little in return. With these patients I don't think he empathized to any degree. But he's interested in psychiatry and accepts responsibility well."

With his neurological supervisors Dr. Abbott got along much less well. One of them described him as "rather an explosive fellow, in terms of judgment—impulsive. I think he's very immature, not very flexible." Another said that he seemed markedly disturbed

emotionally and "was defiant, tended to hide his cases, made a point of trying to wait until the last possible moment and then present the case without giving me the opportunity to give it due consideration." The consultant to the service summed it up as follows: "Dr. Abbott had been exposed to poor indoctrination both in neurology and psychiatry before coming to Winter VA and had a tendency to make snap diagnoses, form impressions on superficial knowledge of the patient, and was somewhat lacking in organization, system and thoroughness in his evaluation of cases. He occasionally tended to go off on tangents, overlooking the obvious to pursue a minor feature of the patient's illness. These negative traits were balanced to some extent by an ability to express himself well, excellent logical thinking when all the facts were in his possession, excellent self-possession before the group, and in most instances a fairly warm relationship with his patients."

During this period, Dr. Abbott had his first experience in a psychotherapy control group. The leader of the group wrote in his evaluation to the School: "The first impression Dr. Abbott makes in a group is that he is a rather clever man who needs to show this by asking provocative and smart questions, but after a few weeks this gives way and Dr. Abbott is then able to take his place with the rest more contentedly. He continues to be sharp and provocative in his comments but is no longer the odd man out. In fact he is then one of the most interested and active members, shows a sensitive understanding of what is going on in the psychotherapy being presented, and is able to make comments about this in such a way as to arouse active discussion but not antagonism. Excellent rating."

In the latter half of his second year, there were no such enthusiastic evaluations of him, and it is clear that his liabilities came more to the fore. He was working on a continued treatment service with chronic, but not severely ill, patients. His service chief wrote the following evaluation after the first two months:

"This resident came to the service with more than average NP experience but does only average work. He works hard, is conscientious and self-sufficient. However, he is limited in his ability to learn from teaching and supervision due to his own defensiveness, which manifests itself by argumentativeness, difficulty in tentatively accepting things which he doesn't understand and a general disagreement with established basic psychiatric postulates. However, the resident apparently does have some capacity to change with persistent and patient handling."

Two months later, the same service chief lowered his rating and reported to the school as follows:

"In the last two months it has been demonstrated that this resident's own personal problems seriously interfere with his evaluation and treatment of too large a percentage of his patient load. There have been several instances of his inability to properly understand the patient and determine a treatment program. He rejects patients who arouse in him personal conflict and this in spite of supervision and getting the resident to become aware of the problem. As the problem repeats itself the resident rejects the supervisor. He continues to be a psychotherapeutic maverick and tenaciously adheres to deviant therapeutic concepts, this being done largely in a defensive way for his own personal needs. Therefore, supervision is difficult and learning minimal. We are in the process of taking up the above matters with Dr. Abbott in an attempt to be constructively critical. It may be we will have to recommend personal therapy but we are not ready to make such a statement now."

Unknown to this supervisor, Dr. Abbott had just terminated treatment after a year, during which time he had had eighty hours of psychotherapy. Several months later, the same service chief continued his story:

"In the last report it was stated that Dr. Abbott's problems were to be presented to him and discussed. This has subsequently been done in some detail. At the time it was felt that if Dr. Abbott continued in the field of psychotherapy he should seek treatment for himself, both for his sake and for that of his patients. Dr. Abbott appeared to respond fairly favorably to the interview and mentioned his need for treatment and outlined some tentative plans which, however, did not include treatment in the immediate future. Instead he made plans to marry which he has since done. In the meantime, his work has improved somewhat and he does what is now passable but far from ideal performance. I sincerely believe Dr. Abbott has tried to understand his difficulties but with no outstanding changes due to new insights. Since his marriage, his attitude has changed somewhat and this reflects itself in his work. However, his work should be closely watched in the future."

Before the time of his marriage, Dr. Abbott was evaluated by the leader of a psychotherapy control group, in which he had for the first time presented a case:

"A difficult case to treat, a borderline schizophrenic. Dr. Abbott started out with the group being rather aggressive and a little defensive. He attempted to overwhelm the group in a way because of his superior knowledge of the patient and the patient's history, but it became very evident as he continued that he has a considerable

amount of anxiety. He is easily aroused by the patient and doesn't have enough objectivity. For the first three months our task was to help him reduce his anxiety, give up his hostility toward the patient, minimize his zeal for treating the patient and stuffing the material down the patient's throat, and on the positive side our job was to show Dr. Abbott that this schizoid individual had a very fragile ego and needed firm support. The patient had a lot of neurotic maneuvers at his disposal and he really led Dr. Abbott around by the nose: in terms of getting him anxious, making him hostile, getting him to act out in small but impulsive ways such as granting a pass one week, refusing it the next week, taking sides with the patient against the nurse on one night and then refusing the patient's medication the next night. After a series of specific advices, Abbott was able to feel apparently that he had some sort of plan or some sort of pat answers at his disposal; was able to reduce his anxieties somewhat and the patient capitulated very quickly. The patient's front of independence and sarcasm disappeared and he became dependent, asking for help. At that point Abbott became very much more comfortable with the patient. The last two or three months, therapy has gone along pretty well. Whenever things become tough for the patient, he reverts back to his neurotic mechanisms but Abbott still gets upset just the way he did. I seriously doubt whether Abbott has really learned much from the experience with the control group that he has been able to consolidate or integrate, although I think it's too early to say that. Abbott gives one the impression of being an eccentric. He wears rather bizarre clothing and unusually long sideburns. He has a shock of black hair that's constantly falling into his eyes and smokes a large, peculiarly shaped pipe so that the net impression is a bit bizarre. He's a man with a lot of tension which is relieved somewhat by mild, not serious or destructive, acting out. I think his chances of making a good psychiatrist are rather poor. He has a certain degree of psychological-mindedness but he is really deficient in ability to empathize with patients."

One of the top administrators in the hospital, in an evaluation interview at the end of Dr. Abbott's second year, said that this resident had come to his attention "increasingly recently because of what are called eccentricities, which seemed to extend into his professional activities. He becomes wildly enthusiastic about new ideas and isn't at all discriminating about them; he has no idea what's good, bad, logical or illogical. At present he's wrapped up in dianetics and has figured out several kinds of off-beat treatment devices. For example, recently he read about a new technique of hypnotherapy; now he cures all his patients with it."

During the summer Dr. Abbott got married. Coincident with this happy turn of events, he was assigned to one of the most prized

clinical assignments, an outpatient clinic. Here he had an opportunity under close supervision to devote himself almost entirely to psychotherapy, something for which he had been eager.

Midway through this final year of residency, the same administrator who had commented on his eccentricities now said:

"I have no personal knowledge of any great change in him, but he was reported six months ago as potentially psychotic; today his supervisors are very enthusiastic—one says even that he is a potential genius! He has been doing mostly psychotherapy and seems to have done a good job. He came in recently and told me that he thought the criticisms I had given him last year had done him good. He had felt that he had always been misjudged, but now he begins to see some of his faults of which he hadn't been much aware before."

A consultant who supervised therapy by Dr. Abbott in his new assignment told us: "I have been very impressed by his sincere attitude to his work, sensitivity, understanding of the patients' problems and his continuing self-searching attitude regarding his own work. His main attributes I would list as interest and sincerity and his eagerness to know things. His main liabilities, I would list as a certain lack of belief in himself which at times makes him dull, but despite it he does very well. He is somewhat dissatisfied with his own work. He has told me he has been experimenting quite a bit with different approaches but perhaps in a somewhat haphazard manner. But I get the impression that there is a real wish to know and learn. I also know him in a group on analytic theory; he has impressed me by very clear theoretical thinking. Deep interest in theoretical problems, some spirit of sound rebellion against authority, and quite constructive." Asked about Dr. Abbott's work with patients, the supervisor described a neurotic woman who "has been handled skillfully and insightfully by Dr. Abbott, who has been very sensitive to this girl's own sensitivity and problems. He has never tried to approach her too bluntly but showed in various ways how he understands her perfectly and follows her, and has been able to inspire in her a great deal of confidence in him, with the result that she began really to express herself for the first time in a very meaningful and quite good way about her problems."

Dr. Abbott was again the presenter in a control group, this time to a supervisor who had known his work and had rated him a good deal lower a year earlier.

According to this supervisor, Dr. Abbott was "quite sensitive to the nuances of patient-doctor relationships. However, my feeling has been in this particular case that with a patient who has been having

considerable difficulty in getting down to the problem for which he seeks help, he has tended to focus too much on his difficulty in getting himself to talk about the problem rather than to take what clues he does present in order to make simple inquiries for further details. In general I have the impression that he is quite a capable man, very sensitive in both good and bad connotations of the word. That is, he's quick to pick up many things about the patient but at the same time is also somewhat narcissistically defensive in his own relationships to authority and perhaps inclined at times with some humor to be unnecessarily argumentative and defensive."

Finally, at the end of his three years in the Menninger School of Psychiatry, he was given top ratings in reports to the School and was described in very complimentary terms in the evaluation interviews for the research project. The supervisor, who had called him "narcissistically defensive," after several more months of contact in the control group mentioned the same quality but added:

"On the other hand, he was able to accept many suggestions and criticisms from the group that were phrased in different ways. I would consider him definitely excellent in comparison to members of his class, but still with need for supervision. He has potentialities for being a rather superior person with further training and supervision."

His chief at the clinic said: "Dr. Abbott has made such great improvements recently that one can consider his work superior. He has demonstrated greatly improved therapeutic and diagnostic skills, handles patient-relationships with less tension and is more objective; he shows now an excellent ability to recognize a patient's real problem (diagnostically) and in his therapy has become more sensitive, adaptable, and is more secure. He always has been very conscientious and eager to learn. Along with this he shows a better personal adjustment, is less eccentric and aggressive. In his relationships he shows more realistic rather than neurotic concern. However, one must recognize that his potentialities at present are limited by certain anxieties and character difficulties. Of the residents who have been here he certainly has shown the greatest amount of growth and improvement." In his evaluation for the selection project the supervisor added that Dr. Abbott had improved throughout his stay at the clinic. "Especially it became possible for him to admit his mistakes. At first it was a battle and he couldn't see any mistakes and he was too subjective. He improved his relationship with staff people, from a lack of concern even for cleaning his fingernails and for dinner table manners to more concern about these. He had good understanding

in relation to the nonpsychiatric staff; he was very conscientious in his notes and entries, there was little difficulty in record keeping. Incidentally, Abbott made good contributions in staff discussions." This supervisor mentioned a few specific cases in which Abbott's assets had come out particularly, and though he tried to think of some cases in which liabilities had showed up, he was unable to do so.

Another supervisor, commenting on the apparently paradoxical fact that Abbott was reluctant to accept criticism or suggestions at face value but really did accept and use them once he understood them, formulated it as follows: "I feel he has a sincere need to *know* rather than to believe."

For the final months of residency Dr. Abbott came back to Winter VA. The manager of the hospital reported in his last evaluation:

"Last night a nurse who has worked with him recently for a few months said he is one of the smartest men she has ever seen. When I asked her about this, she said that he is the only resident she has seen who can present a case quite adequately without any notes."

One of the last evaluations of Dr. Abbott's work for the selection project sums up his whole career as a resident rather well: "I've had Dr. Abbott on this section for only one month at this time. I've known him socially, however, over a three-year period and am most impressed on this present brief supervisory contact with the real change and growth in him. The obstreperousness and contentiousness is largely gone, replaced by a real capacity and willingness to identify himself with and struggle with the problem he faces in the course of his psychotherapeutic work. An excellent thinker, logically organized and at the same time aware of and alert to the emotional aspects of the interaction between himself and the patient. He has real integrity in his capacity to deal very openly with what he sees to be his difficulties and his problems. He maintains his right to differences and is not 'easy' to deal with in terms of quick compliance. From what I've heard about some of the difficulties he has encountered in the school of psychiatry over the past few years, I feel that I am seeing a very different kind and level of functioning. I think in terms of potential capacity he is probably one of the very best of the entire third-year class."

This series of comments forms a pattern in which numerous motifs can be seen to recur, many of them hinted at in the predictive analyses given earlier in this chapter. For the reader's convenience,

Table I-11.1 summarizes the work-relevant points made by the two predictive judges in their final sketches and compares them with supervisors' comments. (The parenthetical numbers in the right-hand column give the approximate frequency with which each point was made by supervisors.) Although not all of the points listed under "Hits" are equally clear and convincing instances of predictive accuracy, the preponderance of Hits over Misses is unmistakable. The Omissions are also impressively few. Under this heading we have listed every quality mentioned by two or more supervisors and not covered in the predictive judges' sketches or their ratings: *E.g.*, "Eagerness to learn and know" was mentioned by four supervisors but is essentially the same as the judges' high ratings for Dr. Abbott on Psychological Curiosity.

Notice, also, that both judges did about equally well despite the fact that in quantitative terms Judge I's predictions were somewhat more accurate. Again and again we have been struck by this fact: Our clinicians could foresee most of the essential qualitative features of a man's work, both his assets and his liabilities, yet could not always predict the quantitative criterion very well. Unfortunately, it would be too enormously laborious and time-consuming a job to make up such an analysis of qualitative accuracy as this for each subject, with the reliability checks that should be included; so we can only give our unsupported opinion that Dr. Abbott's case is not unusual in the extent to which the most salient features of his work were foreseen.\* Judges I and II both saw the two sides of the man—his potentialities for doing quite badly under the pressure of his neurotic problems and for doing good psychiatric work. But Judge II was more concerned with giving the bad with the good and was less optimistic; his predictions were made under the assumption that the subject would *not* get the therapy he needed, although he gave a contingent prediction of what Dr. Abbott would be like if he were treated. His prediction gave a fairly accurate picture of Dr. Abbott's performance during his first two years. Judge I, on the other hand, was more convinced the eventual outcome would be favorable, and his predictions were focused more on the end stage than on the difficulties along the way.

\* It is also interesting to note how much validity there was in both judges' initial impressions. As we shall see in Chap. 12, Judge II's quantitative ratings at this point were so good that his later validities would undoubtedly have been higher if we had not had to analyze the cases "blind."

TABLE I-11.1. COMPARISON OF QUALITATIVE PREDICTIONS  
AND CRITERION DATA

<i>Predictions</i>	HITS	<i>Supervisors' Statements</i>
Hostile to authority (I & II)		Rebellious, antagonistic, defensive to authority (6)
Irritable and at times hypercritical (I)		Argumentative, aggressive (8); hostile toward patients (5)
Unprepossessing front, doesn't always seem sincere, may cause difficulty in management contacts (I); inappropriateness of behavior (II)		Eccentric appearance (3); defensive (3); difficulty relating self to patients (1); peculiar approach to people (1)
Role-playing tendency (II)		Attempted to overwhelm, impress group (2); acts out in mild ways (1); flair for the dramatic (1)
Highly responsive emotionally; undercontrolled but manages to maintain control most of the time (II)		Impulsive, impetuous (4); impatient but became more balanced (2); too responsive to patients (2)
Will have problems of overidentification and countertransference (I)		Overidentifies self with patients (3); not objective (2)
Aware of own difficulties, will overcome overidentification through insight (I)		Recognizes own difficulties; developed insight (5); became more aware and more objective (3)
Serious personal problems that will interfere with his work, anxiety in making decisions (II); anxiety and depression easily stirred up (I)		Emotionally disturbed (5); anxious, tense (5); patients easily arouse or disturb him (4); tends to be depressed (1)
Will turn out well if forms good love relationship (I)		(Work improved markedly after marriage)
Will turn out fairly well if gets therapeutic help (II)		(Work deteriorated somewhat while in treatment, improved afterward)
Has a good deal of warmth for patients, though may seem cool (I)		Warm relationship to patients (3); cold (1)
Can be fairly direct when not emotionally upset (II)		Real integrity, deals openly with his difficulties (2)

*Predictions (Cont.)*

Flair for psychological and psychoanalytic understanding of people (I); psychological sophistication, sensitivity (II)

Very intelligent (I); original, able to be analytic when not too emotionally involved (II)

(Will rank high but) will take a long time to become a really good psychiatrist (I)

*Supervisors' Statements (Cont.)*

Psychologically-minded, sensitive (6); understands patients well, sensitive to nuances of relationship (6) [plus a few other related statements]

Intelligent, brilliant (6); excellent logical thinking (3); constructive, suggests solutions to theoretical problems (1)

[Final evaluations]: Potentialities at present limited by anxiety . . . has shown greatest improvement; potential capacity . . . probably one of the very best

**MISSES**

His empathy is so good that his personal problems won't disturb his understanding of patients (I)

Strikes at others through self-defeating, passive-aggressive maneuvers (I)

Reacts to stress by withdrawal and inhibition of thinking (II)

Strives after great exactness (II)

Deficient in empathy (2); capacity for empathy developing (1)

(no statement of this kind)

(no statement of this kind)

(no statement of this kind)

**OMISSIONS***(Liabilities)*

(No statements about these points, but note he *improved* in both respects)

Intense, undiscriminating enthusiasms (3)

Tends to be dogmatic, inflexible, has preconceived ideas (5)

*(Assets)*

(Possibly reflected in high ratings on Quality of Verbalization, and M. Communicates well)

(Cf. remark quoted from a reference by II)

Handles self well with groups, stimulated group by good contributions (5)

Hard-working, conscientious (4)

[Other assets mentioned are covered by the quantitative ratings and were not repeated verbally by the predictive judges.]

POSTSCRIPT

We hope that by now the reader is interested enough in Dr. Abbott to want to know what has happened to him since he was graduated from the Menninger School of Psychiatry. He has been good enough not only to give his permission to have this case history published but to supply some additional details about his subsequent career.

His subsequent career has been marked by consistent achievement, although its pattern is not very different from that of many another young psychiatrist. He went west again, although not to his home state, and took a job with a government institution. After a couple of years he became its director, not very enthusiastic about administrative work but doing it adequately, supervising psychotherapy and doing a lot of it himself. He was admitted to psychoanalytic training and was just finishing his training analysis at the time this chapter was written. He and Mrs. Abbott have two children and appear to be happily married.

Both authors had an opportunity to see Dr. Abbott after he had read this chapter and were impressed with the diminution of insecurity and his growth of self-confidence and maturity. He still retains a humorous twinkle and an informal easiness of manner; the tension that was so noticeable when we first met him is less prominent. There is still indication that he will continue to grow as a psychiatrist and live up to the promise the predictors saw in him.

# 12

## PRINCIPAL RESULTS OF THE SECOND DESIGN

A SINGLE EXAMPLE, SUCH AS THAT OF DR. ABBOTT, GIVES A RICHER AND more meaningful account of successes and failures in Design II predictions than statistical results, although, perhaps, a more misleading one. No one instance can be truly typical; what appear to be "Hits" and "Misses" in predicting cannot be taken at face value as long as there is the possibility that they might have been coincidences. Since statistics were invented for the purpose of making sure investigators do not fool themselves in this way, this chapter summarizes the main results of a statistical analysis of the results of Design II. The analysis itself is set forth in the technical supplements to this chapter (Appendices 12.1-12.9).

### CROSS-VALIDATING THE MANUALS

Design II was undertaken in an effort to improve on the job of selection being done. The design we worked out and tested in our final Predictive Study with Classes V and VI was a complex one; it had three main objectives in addition to the one just mentioned:

1. To increase the objectivity and the precision of methods used to analyze the rich data from projective tests and interviews

and to compare the "objectified" predictions with free predictive ratings;

2. To compare predictions based on varying amounts and types of data, in the hope of finding an optimal battery of procedures;
3. To find the level of generality or specificity on which the most valid predictions about psychiatric work could be made.

The manuals guiding analysis of the TAT, Rorschach and other tests, and the Interview were designed because of the first numbered objective. Some data from the first trial of several manuals have already been briefly cited; we shall now examine in more detail how well these manuals worked in the final Predictive Study.

#### THE ATTEMPT TO MAKE CLINICAL JUDGMENT OBJECTIVE

We wrote the manuals in a dual attempt: to focus clinical instruments on the predicting of psychiatric competence; and to make the process of analyzing complex clinical data explicit enough so that different people could go over the same material and come out with the same prediction. If we failed in this second respect, then it would be difficult to demonstrate that the manuals had been very helpful, even if individual judges were able to get good results with them. Valid predictions almost always require, first of all, that a method be independent of the special skills or blind spots of individual users. Nevertheless, we were interested in agreement between scorers of the manuals only secondarily, being convinced that a technique may be useful even if it is applied differently to the same group of subjects by different clinicians, as long as they all can reach useful conclusions.

Each manual listed signs, or *cues*, thought to be prognostic of good or poor psychiatric performance. The algebraic total of scores on these cues (*i.e.*, the number of unfavorable signs subtracted from the number of favorable ones) is called the "cue-sum," for short; it is the main predictor yielded by a manual. The statistic that tells most about "observer agreement"—the tendency of judges to score alike—is, then, the correlation between the cue-sums of different judges, using the same manual on the same subjects.

In terms of this measure, the TAT Content manual was quite acceptably objective; the average correlation between the three judges who used it was over .8. This figure compares well with those achieved

with most other scoring systems for the TAT (see Tomkins, 1947, p. 4). The Picture Reaction Test manual gave a correlation almost as high (.77); cue-sums of the TAT Self-Interpretation and the Rorschach manuals had slightly less observer reliability (correlations for both being about .7) but were still at a fairly respectable and usable level.

Two of the manuals fell markedly below the level of objectivity of the others. Cue-sums of three judges using the TAT Formal Aspects manual tended to correlate with each other on the average only about .3. The Interview manual was virtually worthless in standardizing the analysis of this particularly elusive method: cue-sums of two judges correlated no better than chance (less than .2). Similarly, when the figures are examined for agreement on scoring individual cues, many of them (even in the TAT Content manual) were used quite differently by different judges (see Appendix 12.1).

These figures on scorer agreement do not constitute either a uniform success or failure. The two manuals that called for the greatest amount of judgment (Interview and TAT Formal) were found most wanting, leaving plenty of room for different judges to get various results. If the most objective manuals had good validities for one judge, however, there would be more likelihood that they would for another as well.

#### VALIDITY OF TAT MANUALS

The first cross-validation of the TAT manuals was most encouraging (see Appendix 10.7), and we felt that our revisions after an inspection of these results with Class IV had made the manuals even more likely to shine in the final test, with Classes V and VI.

The Content manual did not show useful validity on the second cross validation. All three of the judges' cue-sums failed to predict either Supervisors' Evaluations or Peers' Evaluations of any aspect of psychiatric work better than chance. (Validity correlations are mostly about zero, ranging from -.22 to .21.)

After the first cross-validation, we had selected a cutting-score for each manual (see Appendix 10.7), such that every candidate scoring as low as the cutting-score or lower would be rejected, everyone above would be accepted. Obviously this must be done in such a way as to maximize the numbers of Inadequate residents and Drop-outs (see discussion of the Qualitative Criterion, Chap. 8) who would be rejected and to minimize the number of Adequate residents sacri-

ficed—particularly Superior ones. When this cutting-score for the Content manual was applied to Classes V and VI, it failed to separate acceptable from unacceptable residents better than chance, no matter which judge's scoring was used. (Details of the validation of all manuals are given in Appendix 12.2.)

We had one afterthought about TAT Content which did not get fully tested, yet which looks promising enough to be reported. That was a rating of the *adequacy of the hero* of each story to attain his goals. It had originally been tried out in the early stages of the manual for Formal Aspects; it was removed from that manual because it seemed to be a matter of content and was overlooked until we had already scored several cases of the predictive study. The decision was then made to score it systematically, even though there was no descriptive and defining material in a manual to guide us in doing so. The reliability coefficient for the 47 cases scored by Judges I and II was .78, which is good for so judgmental a variable. The algebraic sum over the 12 stories for each subject (clearly adequate behavior of the principal figure being scored plus one, clearly inadequate behavior minus one) was a more valid predictor of most criteria than all the rest of the TAT manuals. It correlated .28 or better for both judges with Peers' Evaluations of Administrative Competence and showed distinct promise (correlations for both of .20 or more) as a predictor of Diagnostic and Psychotherapeutic Competence (see Appendix 12.2 for further details).

The less objective Formal Aspects manual also showed promise. One judge was unable to do better than chance with it, but for the other two—especially the one who had written it and presumably understood it best—the cue-sum correlated better than chance with some of the in-training criteria. For this judge it correlated between .22 and .28 with Over-all Competence and Diagnostic Competence (either Supervisors' or Peers' Evaluations), and for both of them the cue-sum correlated over .3 with Spontaneity *vs.* Inhibition.

The cutting-score from the Formal manual (as scored by its author only) would have worked rather well as a dichotomous predictor: it would have resulted in less than 5 per cent of Extreme Errors (see Appendix 9.5), when simply accepting everyone would have meant 19 per cent of Extreme Errors—a finding significant at the .02 point. This judge would thus have correctly turned down 10 of the 13 unacceptables, although at the high cost of 13 Adequate and 3 Superior residents.

Parenthetically, results from a later, independent attempt to predict competence of labor mediators proved that parts of the Formal Aspects manual had a good deal of validity (Landsberger, 1956). For example, the Stereotypy of Stories cue correlated —.5 with Originality of Ideas, as rated by criterion judges, while the Genuineness cue correlated about .5 with three other criterion variables. It seems that the Formal manual contains valid measures of a number of aspects of personality and may be more generally useful if the manual can be made explicit and clear enough so that reasonably well-trained clinicians can use it. (See also Holt, 1958b.)

Returning to Classes V and VI, the manual for the Self-Interpretation of the TAT performed much like the Formal manual in that one judge's cue-sums predicted nothing, while the other two correlated better than .3 with one criterion, in this instance Peers' Evaluations of Diagnostic Competence. (In addition, the cue-sum of the manual's author was also a significant predictor of three other criteria, including both Peers' and Supervisors' Evaluations of Psychotherapeutic Competence.)

The changes in procedure and manual between the application of the Self-Interpretation to Class IV and to the Predictive Study classes make it a little dubious to view this as a strict cross-validation. The doubtfully applicable cutting-score derived from Class IV did not seem useful as a selection device (see Table 12.8).

Considering the many factors tending to hold down validity—a restricted range of talent in the sample (Menninger School of Psychiatry residents only), brief scales (the Self-Interpretation manual had only twenty-five cues), less than perfect reliability of the criteria, and testing errors of various kinds—the record of the TAT Formal and Self-Interpretation manuals on cross-validation is good. These results suggest that it would be worthwhile for persons who are engaged in psychiatric selection to devote further effort to these manuals, bringing out their potentialities for greater objectivity and greater validity as predictors.

*Rorschach*.—The manual worked out by Dr. Walter Kass for scoring this most widely used and respected of projective techniques had been subjected to extensive pretest and revision. In its final trial, however, it made about the same kind of record as the TAT Formal and Self-Interpretation manuals: the manual's author was able to get a couple of validities of about .25, and one other judge's cue-sum also correlated over .2 with Diagnostic Competence (Supervisors'

Evaluations). The third judge's cue-sum correlated consistently near zero with all criteria, and the general level of validities was low for the other two (see Table 12.10 in Appendix 12.2).

When we analyze the Rorschach results in terms of the cutting-score and its effectiveness, we should recall that this manual was planned in a slightly different way from the others. The few cues that it contained were all "stop items," the presence of any one of which in a Rorschach was to be a warning signal calling for a more careful evaluation. In one way, this cutting-score worked very well: Accepting without question everyone who received no black mark by his Rorschach performance would have let into the School no Inadequates and one or two Drop-outs, by the scoring of Judges IV and I respectively, and not one of either category by the scoring of Judge V. (The results are statistically significant for Judge V only; for Judge IV,  $p = .06$ .) It is questionable whether such a fine-meshed screen would be useful, however, since it caught about half of the acceptable residents, no matter which judge was wielding it. If the Rorschach had been used in this manner, it would have been necessary to assess carefully by other techniques a very large proportion of the applicants. (Lowering the cutting-score for the Rorschach manual, on the assumption that several stop-items must mean that a man is really unpromising, would not have increased its usefulness.)

*Picture Reaction Test.*—There is little to say about the PRT manual except to report briefly that it failed completely to yield any evidence of validity. Promising-looking differentiations in the Small Sample regressed quickly to chance results, regardless of the judge, and whether the data are analyzed by means of correlations or by cutting-scores.

*Interview.*—The Interview manual's cue-sum correlated surprisingly well with Peers' Evaluations. In both judges' scoring, it correlated between .2 and .3 with Over-all Competence, Psychotherapeutic Competence, and Diagnostic Competence. As usual, it worked best in the hands of its author, whose cue-sum also predicted several other criteria at the same level of significance. Except for the fact that it did not work as a predictor for the other judge (except against Peers' Evaluations), the cue-sum of the structured Interview of Design II was about as effective as the judgments of most interviewers in Design I.

There was no prior trial of the Interview manual in the form used in the Predictive Study (Design II), so no cutting-score was

available for cross-validation. It is possible to find places in the distribution of scores by either judge that distinguish Acceptables from Unacceptables at the 5 per cent level or better, although the cutting-scores are very different for the two judges (see Appendix 12.2).

To conclude the discussion of the manuals method: There were some modest successes as well as some resounding failures of this method. Several of the manuals yielded scores that could have been used in actual selection with fairly good results, provided that the right judge was scoring them. As they stand, however, these manuals contain enough ambiguity and require the exercise of enough unguided judgment for there to be no way of guaranteeing that anyone other than a manual's author would be able to get results with it.

A great deal more work would have to be done, therefore, before one could consider any of our manuals a finished and trustworthy instrument. Item analyses were done for all manuals as a final step, and they revealed a core of items that consistently discriminated between better and poorer residents, from Small Sample through one or two cross-validations and in the scoring of more than one judge. Doubtless this fact means that if sufficient energy were applied to the task of making up new cues, rewriting and reworking them until they were reliably scored and held up under successive cross-validations, instruments with practical usefulness could be developed. The job would, however, take further years of work with large populations.

The manuals method was a compromise between unguided intuition and atomistic oversimplicity. We undertook it with hopes of retaining, but systematizing, the kind of global and interactional approach we were using in our daily clinical work with the instruments. Yet the closer we clung to the clinical method and categories that demanded clinical skill, the more difficulty we had with scorer agreement (as with Interview and TAT Formal Aspects), and the further over toward "objective" itemization we swung (TAT Content, PRT), the less valid the result. The only exception to this rule—Adequacy of Heroes in the TAT—deserves a thorough trial by other workers.

#### PREDICTION BY FREE RATINGS: MAJOR VARIABLES

The second principal objective of the predictive study was to see how well our judges could do in forecasting the performance of

residents in the School by means of free ratings (like the predictive ratings of Design I), based on varying amounts of material. Free ratings are contrasted with the rigid cue-sums of the manuals method, in that the judge was encouraged to use all of his clinical acumen, any hunches he might have about a case from the unique configuration of the data, and his accumulated experience, whether or not it had been codified or perhaps even verbalized.

(Table I-12.1 presents a sample of the validities of these free ratings. It is a section of the larger Table 12.13 in Appendix 12.3, which contains the same information on all five major aspects of psychiatric work. This table is enough to make the essential points, however.)

First, most of the correlations are quite low. In the total table, they range from  $-.16$  to  $+.58$ , only half of them being over  $+.15$ . At the final stage, however, when the two judges had access to all predictive data, the validities are remarkably good. It appears that clinical judgments based on blind analysis of a recorded interview, a single projective test, or any combination of these that we tried short of the total file of material, were no better than the unsystematic judgments of the rule-of-thumb method (Design I). Only when the judges integrated interview material, projective tests, nonprojective tests of intelligence and association, and background data into a total understanding of the case were they able to attain a level of predictive efficiency notably above anything previously obtained in this research. Nevertheless, even in the earlier stages of analysis, a judge usually predicted more successfully as he got more data to work with.

In general, the predictions correlated more highly with the criterion judgments given by the residents of each other than they did with Supervisors' Evaluations. This tended to be true no matter what the aspect of work under consideration, and for all judges except I (and V—see below), with a few minor reversals. There was no such tendency on the part of psychiatric interviewers; in general their predictions (on this and earlier classes) tended to correlate more highly with Supervisors' Evaluations. We have no very plausible explanatory hypothesis to offer, but it is possible that the results reflect patterns of identification with either the supervisory or the student groups (see, however, Appendix 12.4).

Even though the correlations on the whole are low, still there are consistent trends of difference between the judges. Judge I's validities tended to be higher than those of the other judges, though not on all variables at all stages of analysis. As we shall see later, however

(Table I-12.2), in terms of two other measures, Judges I and II did equally well at the Final Stage.

TABLE I-12.1. VALIDITIES OF PREDICTIVE RATINGS OF OVER-ALL COMPETENCE, DESIGN II (CLASSES V & VI)

BASIS OF PREDICTIVE RATING	SUPERVISORS' EVALUATIONS <i>Predicted by Judge:</i>				PEERS' EVALUATIONS <i>Predicted by Judge:</i>			
	I	II	III	IV	I	II	III	IV
Stage 0 (Impression)	.12	-.05	.12		.27	.26	.18	
Stage 1	.26	-.10	.14	.11	.23	-.02	.17	.02
Stage 2	.26	-.08	.08	.12	.21	.09	.13	.04
Stage 3	.28	.03	.07	.15	.25	.22	.11	.17
Stage 4	.29				.27			
All Data	.57 <sup>a</sup>	.22			.52 <sup>b</sup>	.48		

<sup>a</sup> N = 38.

<sup>b</sup> N = 31.

NOTE: Blanks indicate that ratings were not made for the judge-stage combination in question. Numbers of cases vary from 46 to 64 in the left half of the table, from 43 to 45 in the right half, unless otherwise specified. Different data were available to the judges at any one stage; see Chap. 10.

To read this table, for example, look at the number occupying the second row, fourth column, of the right half of the table: it reads .02. This means that one series of predictive ratings correlated .02 (*i.e.*, not at all) with one of the criterion measures. The predictor was Judge IV's rating of Over-all Competence at the first stage of analysis, after she had seen only the Rorschach test. The criterion was Peers' Evaluations of Over-all Competence.

Differences in the predictability of various types of work were slight and inconsistent, though, in general, Peers' Evaluations of Diagnostic Competence were predicted consistently best by all judges at all stages, and Supervisors' Evaluations of Psychotherapeutic Competence were predicted least well (see Appendix 12.4).

Correlation coefficients such as those in Table I-12.1 are less meaningful than cutting-score analyses. When we look at the results in terms of what the practical consequences would have been if our Predictive Study judges had actually been selecting residents by means of their free ratings, the success of Design II becomes even clearer (see Table 12.15). At the final stage, both Judge I and Judge II made the basic distinction between Satisfactory and Unsatisfactory applicants—which is the essential aim of selection methods—with statistically significant accuracy. Judge I gave “Accept” ratings to only two of the five Unsatisfactory subjects whom he rated at that

stage, and "Reject" ratings to only two of the eighteen Adequates and to no Superior candidate. Judge II would have eliminated half of the six Drop-outs at the cost of only one Adequate resident. Both made significantly fewer extreme errors than would be expected by chance ( $p < .025$ ).

Surprisingly enough, Judge III's rating of Over-all Competence at Stage 1 (Interview only) was the only other rating to have a significant degree of validity. But he was very cautious about giving ratings in the reject range: he rated only three cases so low, one of whom turned out to be Borderline, one Inadequate and one Drop-out. This constitutes one of the few exceptions to our general conclusion that single sources of data were not very useful. It seems that Judge III tended to be correct when he decided that a man was not acceptable but could not predict degrees of competence, and he was unable to make good use of additional information about a case (TAT and Self-Interpretation).

#### PREDICTING QUALITIES OF RELATIONSHIP WITH PATIENTS

It will be remembered that predictive judges in Design II were asked to estimate the probable performance of psychiatric residents on fourteen "minor variables," mainly representing types of relationship with patients. (Table 12.16, in Appendix 12.5 presents all 154 correlation coefficients along with some commentary for anyone who is curious about the details.) The average correlation is low, but some of the validities are quite good—many more than could have occurred on a chance basis. Let us see then what distinguishes good validities from poor ones.

Certain qualities were much better predicted than others. *Judiciousness vs. Impulsiveness* was the most predictable, especially by Judge II, whose predictions correlated over .3 at each stage, going up to .55 at Stage 3 (when he had seen the Interview, TAT and Self-Interpretation). Judges I and IV predicted this variable fairly well also, but Judge III drew a complete blank; using the same data at Stage 3 as Judge II, his predictions correlated only .06 with the criterion. Nevertheless, a majority of the attempts to predict this variable succeeded, and the same can be said for two others: *Spontaneity vs. Inhibition* and *Acceptance of Responsibility* (validities up to .33 and .30 respectively). At the other extreme, 6 of the 11 validity coefficients involving *Overidentification* were negative, and the distribution ( $- .12$  to  $+ .16$ ) was just about what one would expect from chance

alone. There is no obvious characteristic that distinguishes the more predictable variables from the less.

As in the case of the major variables, judges differed in ability to forecast these aspects of psychiatric performance, but much less consistently. When the judges were given additional information, there was a trend toward increased validities here too, but less markedly than in the case of the major variables.

#### HOW DOES IT AFFECT THE VALIDITY OF PREDICTIONS TO GIVE JUDGES MORE DATA?

One of the chief aims of the Predictive Study was to test the effectiveness of predictions based on varying amounts of data. If it had not been for the addition of the final stage of analysis for two judges, we should have had somewhat inconclusive findings. It is true, the average correlation over all judges and all major variables goes up steadily from Stage 1 to Stage 3, and the number of significant validities increases after Stage 2; but if we look at the data in terms of the individual judges, the results are less clear. Judge II's validities rose somewhat throughout, but Judge III's showed an unmistakable falling trend as he received more data about each case. The other two judges usually increased in accuracy from Stage 1 to Stage 3, but not markedly. The final, All-Data stage produced a very pronounced advance in validity for Judges I and II, however.

From one standpoint, the comparison of the free rating and the cue-sum also tells about the effect of adding more data. The manuals guided attention to signs intended to indicate good and poor psychiatric potential, but they used only a part of the information contained in the test or interview. Part of the objection to projective techniques and interviews is that they contain too much and put too severe a strain on the capacity of clinicians to select what is meaningful and disregard what contributes only error. Did our manuals succeed in keeping us on the strait and narrow?

The evidence points again toward differences between judges in integrative capacities. For Judges II and IV, using a single test, there is no difference between their validities, whether only the cue-sum is the predictor or whether they could use every clue they could find in it. Judge I was able to do significantly better with the PRT when he made a free rating (*Over-all Competence vs. Supervisors' Evaluations, r = .26*) than when he stuck to the cue-sum ( $r = .00$ ).

And Judge III did *worse*; his cue-sums from the Interview correlated significantly with both Supervisors' and Peers' Evaluations of Overall Competence (.23 and .27), whereas his predictions of that variable had insignificant validity (.14 and .17). The interesting thing about these results is that Judge I was best able to profit by adding more information at later stages, and Judge III was the only one whose validities tended to drop off as he added test data in subsequent stages of analysis. Two cases do not constitute a very firm finding, however, especially since the data have an insignificant contradictory trend in the case of one other judge.

At any rate, we can conclude tentatively from our data that the effect on predictive validities of adding information, in either large amounts or small, depends a great deal on the person doing the predicting. In a study with a total of only five people making predictions with varying amounts of material, we can hardly generalize any further.

In other researches, too, both trends have been reported, though the bulk of the evidence leans toward greater validities with increasing amounts of information (see Appendix 12.6, where this topic is treated at greater length). But it is almost impossible to study the effect of increasing information *alone*, without varying the relevance of the data to the kind of behavior being predicted and, indeed, without changing the very nature of the predictive process. If a forecaster is to make use of increasing amounts of information, he must be able to select the relevant and ignore the irrelevant or distracting, to integrate and organize large amounts of data without getting confused or overwhelmed. It has recently been shown on both logical and empirical grounds (Cronbach, 1955; Meehl and Rosen, 1955) that added valid information is likely to lead to poorer results than using some single fact, such as the probable frequency of failures, unless employed with great care and precision.

Probably there are two types of predictive inference in these studies. The first is a relatively simple matter of extrapolating from past performance. When one has a set of letters of recommendation, telling the kind of work a man has done in his internship, and transcripts showing the type of work he has done in medical school, together with other types of data directly relevant to his performance in tasks that have some degree of similarity to the work that is being predicted, no very high order of inferential wizardry is necessary to make a prediction with some modest chance of validity.

The other type of inference is much more laborious and indirect. Essentially the judge constructs a model of a personality and deduces the appropriate kind of behavior from that model. This means that he must take data of whatever kind he can get and put together a total picture of a personality that would have produced such data. He starts with personological constructs that do not relate in any very direct way to work, inferring from them to constructs that do bear directly on the performance concerned, and, finally, attempting to predict that performance.\*

Clearly, in the second and more elaborate type of inference there are many more places at which the judge may make errors; the inferential chain is longer and involves many more assumptions. It is important to note that the apparently simple issue of the amount of information is thus confounded with a shift from the simpler to the more complex type of inference in many studies.

In summary, adding information may have various effects on the accuracy of forecasting behavior, depending on: (1) the kind of data added, in relation to (2) what is being predicted, (3) the integrative capacities of the judges, (4) the kind of inferential approach followed by them, and probably (5) the order in which the data are presented—all of these factors interacting with each other. There can be no one answer, therefore, and it should be possible to choose combinations of these five classes of determinants to give almost any kind of results.

#### ON WHAT LEVEL OF GENERALITY ARE THE MOST VALID PREDICTIONS MADE?

Much of the complexity of the Predictive Study, the culmination of Design II, was due to our desire to investigate the issue of the optimal level of specificity at which to make predictions. Perhaps the results in Design I on predicting degrees of Over-all Competence in psychiatry were not better because this was too global a kind of performance to be predictable. Therefore, we used variables on four levels of specificity: (1) Over-all Competence, the most general; (2) Competence in Psychotherapy, Diagnosis, and the other major variables; (3) the fourteen qualities of relationships with patients; and

\* In his recent book Meehl has made a similar analysis of the process of clinical prediction (Meehl, 1954, Chap. 6).

(4) most specific, the sixty-four relatively concrete statements about psychiatric work.

We found no clear-cut differences. There is simply no convincing evidence that predictions on one level of generality consistently differed in validity from predictions on another level (see Appendix 12.7). For the most part, this negative result may be due to the fact that our various predictions and our criterion ratings of different aspects of work called for a great deal more differentiation than either set of judges could make. A statistical study of the correlations between the criterion ratings themselves found them to be minor variations on one main over-all impression of how good a man a resident was; and the same was true of the predictive ratings (see discussion of the factor analyses of predictors and criteria, Chaps. 8 and 14). Under such circumstances, one would expect a prediction of one variable (say, Psychotherapeutic Competence) to correlate about as well with one criterion rating as another—*e.g.*, it might even correlate better with Management Competence, than with therapeutic competence. That is exactly what happened; with few exceptions, the predictors showed no consistent tendency to correlate best with the appropriate criterion. (Details are given in Appendix 12.8.) Therefore, if a “specific” variable turns out to be highly correlated with all other “specific” and “general” ones, there is little use in asking which are best predicted.

It is regrettable that better criteria of specific types of behavior were not available and that we are left with such an unsatisfactory concept as Over-all Competence. True, for practical purposes, this is enough; a man can be quickly accepted or rejected if one considers only whether or not he meets minimal standards on a single general dimension. Theoretically, however, this state of affairs is more troublesome. The approach to clinical prediction that makes most sense to us begins with the proposition that we make primary inferences from information about a man to a relatively specific level of behavioral tendencies, or intervening variables, such as our thirty-two variables of personality. Then from there, we make secondary inferences about various reactions of a subject in given situations. Ideally, then, such an evaluation as Over-all Competence (which is not itself a form of directly observable behavior) can be predicted only by a complex chain of inferences involving at least three links—and, consequently, more opportunity for error than in predictions

that use only primary inferences. Perhaps this analysis is overelaborate, and clinical predictions are more frequently direct hunches; our data allow no further clarification of the issue.

#### A COMPARISON OF DESIGNS I AND II

Now that we have seen the main results of Design II, let us review the findings from the earlier design and see what may be learned from comparing the two. This comparison is hampered from the beginning by the fact that it has to be limited to in-training criteria, whereas the most important and statistically reliable results of the original study were the successful prediction of establishing psychiatric practice and passing certification examinations (see Chap. 9). But at the time of writing, the residents of Classes V and VI had not been graduated long enough to have had an adequate opportunity to become diplomates of the American Board, and the mail follow-up had been terminated even earlier.

In terms of in-training criteria, most of the validities achieved by the interviewers and testers in Design I were positive, but low. For all classes, the average rating by psychiatric interviewers correlated .24 with Supervisors' Evaluations of Over-all Competence; the corresponding figure for the School's psychological testers was .27. These levels of validity were as good as the results usually obtained in vocational selection studies, but no better.

The present chapter has made one conclusion plain enough: Design II brought forth little that surpassed the earlier work. The recorded, structured Interview gave results no better than most of the School's psychiatrists were achieving and worse than the best of them could do. Our manuals did not turn single projective tests into objective instruments that could do a better job of predicting than the clinical synthesis of several tests by the psychologists on the Admissions Committees. What remains is the performance of the authors with the maximum of information: Judges I and II at the All-Data stage of analysis.

In Table I-12.2, we have brought together figures on validity from both designs, keeping the sample of subjects as much the same as possible (Classes V and VI).

This group of applicants was a somewhat easier one for the School's interviewers and testers to predict about than most of the classes they had worked with but was still not the one on which they

TABLE I-12.2. COMPARISON OF THE VALIDITIES ACHIEVED IN DESIGNS I AND II (ON SUBJECTS FROM CLASSES V &amp; VI)

CRITERIA	PREDICTORS: OVER-ALL RATINGS				
	DESIGN I		DESIGN II		
	Inter-viewers' Mean Rating	Psycho- logical Test Rating	All data Jdg. I	Jdg. II	Rorschach Jdg. V
<i>Supervisors' Evaluations of:</i>					
Over-all Competence	.29	.39	.57	.22	.07
Psychotherapeutic Competence	.23	.36	.48	.15	.01
Diagnostic Competence	.37	.33	.58	.24	.24
Management Competence	.25	.40	.52	.13	.09
Administrative Competence	.41	.35	.55	.24	.11
<i>Peers' Evaluations of:</i>					
Over-all Competence	.25	.47	.52	.48	-.04
Psychotherapeutic Competence	.26	.50	.55	.36	.00
Diagnostic Competence	.31	.46	.42	.42	-.05
Management Competence	.15	.34	.36	.24	-.06
Administrative Competence	.19	.32	.42	.27	.09
Per cent Extreme Errors <sup>a</sup>	18	6	3	8	13
P-value for discrimination of Satisfactory from Unsatisfac- tory subjects	>.20	.02	.02	.03	>.20
Number of subjects	64-68	64-65	30-39	64-68	45-50

<sup>a</sup> See Appendix 9.5 for explanation of this measure. Italicized numbers indicate significance at the .025 point.

did best. Nevertheless, their validities stand up rather well in comparison to those obtained by the best predictive judges of Design II. Judge I's validities look distinctly higher, but the comparison is a little suspect since there is no way of knowing whether he would have been able to maintain the same level if he could have rated all cases.\* Besides, in statistical terms, his validities are not significantly higher than those of the School's tester, who used a single over-all prediction.

The latter's performance is quite impressive when one realizes that he was working with just the Wechsler-Bellevue, Rorschach, Word Association and Szondi tests—plus his direct impressions of the candidate. Judges I and II, at the final stage, had all of these tests (except the Szondi to which the School's judges paid virtually no

\* He also may have been somewhat more "contaminated" than the others; see Appendix 12.9.

attention anyway), *plus* Interview, TAT, Self-Interpretation, PRT, Credentials, and Supplementary Face Sheet! But the Predictive Study judges were working blind and may have lost as much in that way as they gained by their many sources of information about the candidates. The fact that validities based on Initial Impressions alone were, for most judges and most criteria, at least as good as those based on a careful analysis of the first stage, and in some instances were dramatically better, also suggests that a good deal of validity was lost by the necessity for blind analysis. Judge II seems to have been particularly penalized by the necessity to work blind; he was able to attain the surprisingly high validity of .39 (Psychotherapeutic Competence, Peers' Evaluations) from his impressions of the candidates alone, a figure he was unable to surpass after a blind analysis of all the test and interview data.

Incidentally, the distribution of validity correlations for initial impressions (from — .10 to + .39) is very similar to the validities obtained by the School's psychiatric interviewers, who were operating in much the same way—making a rating immediately after an hour of clinical contact, without having the opportunity to study the data elicited in that time and without the putative assistance of manuals. When combined, however, these ratings did not average out to form as good a predictor as the psychiatric interviewers' ratings: the validity of the mean Initial Impression of Over-all Competence for Judges I, II, and III was only .10 (against Supervisors' Evaluations).

It happens that thirty-six of the cases were tested and evaluated by Judge V, serving in the role of the School's tester. Later, going over the Rorschachs of the same men (and enough others to make a total of forty-five), he attempted to make predictive ratings from this technique alone, with the identities of the cases concealed. The right-hand column of Table 12.2 gives the meager result. This judge, with a single over-all rating based on three tests (which formed the core of his standard clinical battery), could get validities against criteria he was not even thinking about that were as good as those of the Predictive Study judges who had access to all data and were aiming their predictions specifically. Yet without the tests of intelligence and association, and without his clinical impressions, Judge V was helpless.\*

\* Similar negative results were obtained at the Postdoctoral Institute for Psychotherapy in New York City, where the attempt was made to predict the adequacy of performance of psychotherapists in training, using the Rorschach alone. Validity coefficients were near zero (Abel, Oppenheim & Sager, 1956).

This comparison helps to underscore a couple of points already made: the ability of the judge is important, but he must have an adequate body of data to work with. We did not discover any brief battery of procedures that was better than the one used by the Admissions Committee's tester, but we did establish that a single projective test was *not* adequate for this kind of prediction. Even though the Rorschach or TAT cannot be used alone, they probably do contribute usefully in the context of a rounded body of data. Psychiatric residents can be selected fairly well by a single interview, but it is difficult to choose the unusual interviewer who is able to do it.

To summarize: The method of guided clinical judgment, as we used it here, meant studying the performance to be predicted, surveying the personality variables that must be assessed, making a preliminary application of various tests and interviews to residents of known degrees of competence, systematically organizing test and interview evidence in manuals to facilitate the particular predictions to be made, and carrying through preliminary cross-validations of all these procedures. Its main effect was apparently to raise the ceiling of the predictive accuracy a judge could attain, without guaranteeing that any particular clinician would be able to reach it. All this elaborate method, as compared to the direct application of common sense and clinical experience, succeeded so far as it showed that quite an impressively accurate job *could* be done with a complete body of information about an applicant. We had hoped that our manuals would provide enough external guidance so that any reasonably competent clinician could be assured of attaining at least a minimally significant validity of prediction; in this we did not succeed. The failure was probably due to two principal factors: First, it is doubtful that a single interview or test, no matter how meticulously analyzed, gives a broad enough sample of a physician's behavior to predict from. Second, the process of framing cues (items of behavior with some determinate significance), validating them, rewriting and cross-validating them could not be pushed far enough in the time available to us, though we still believe that the method is a promising one.

Perhaps the briefest way to summarize this complex chapter is to say that there are no short-cuts to highly efficient prediction about such complicated human behavior as effective work in a psychiatric residency. Long experience, careful study of the problem, talent of some unspecified kind, and enough data to work with—some combination of these is necessary.

## VALIDITIES OF "LIKING" AND SOME ADDITIONAL PREDICTORS

IN THE PREVIOUS CHAPTER, WE PRESENTED ALL OF THE VALIDITIES OF cue-sums from the manuals and the validities of free clinical ratings. They are the central findings of the research; yet quite a number of other predictors was tried out, and a number of interesting points remains to be made in connection with them. For one thing, we have yet to present our highest correlations between predictors and criteria.

### "LIKING" AS A PREDICTOR

The ratings made by predictive judges on Classes V and VI included one variable that was not intended as a predictor but as a control: how well the judge liked the person whose productions he was analyzing. We knew that there was a danger of the so-called "halo effect" in evaluative ratings and thought that if we were to record a separate rating of how well we liked a candidate, this would help us to make independent predictive ratings free from naive bias in favor of people who were personally appealing. For similar reasons, we asked for ratings of personal liking when we obtained criterion ratings from supervisors and from the residents themselves.

In an early exploration of the findings with preliminary criteria, we thought that we would try to separate out the effect of our Liking ratings by the statistical method of partial correlation. A preliminary step in calculating such a coefficient is to correlate Liking with the criterion as well as with the predictor. What was our astonishment to find that the Liking rating seemed to have more "validity"—*i.e.*, it correlated more highly with the criterion—than the predictive rating itself! Consequently, we looked into other such possible correlations with final criteria.

The results are contained in Table I-13.1 (a sampling of the complete set of validities against all criteria given in Tables 13.2 and 13.3 in Appendix 13.1). Since Liking was rated at each stage of analysis by every judge, these tables have been set up closely following the form of the tables in Chapter 12 where the validities of the predictive ratings themselves are presented; *e.g.*, comparing Table I-13.1 with Table I-12.1, we can see a great deal of similarity. There is the same difference between judges, the same superiority of the validities against Peers' Evaluations over those against Supervisors' Evaluations, and similar idiosyncratic patterns of rising and falling validities (depending on the particular judge, with increases in amounts of data at the different stages of analysis). Moreover, there were again differences between the predictability of the various aspects of psychiatric work (Tables 13.2 and 13.3). The main difference is that the coefficients in Table I-13.1 are more often positive and highly significant than those in Table I-12.1!

TABLE I-13.1. VALIDITIES OF LIKING RATINGS AGAINST OVER-ALL  
COMPETENCE, DESIGN II (CLASSES V AND VI)

LIKING BASED ON:	SUPERVISORS' EVALUATIONS <i>Predicted by Judge:</i>				PEERS' EVALUATIONS <i>Predicted by Judge:</i>			
	I	II	III	IV	I	II	III	IV
Stage 1	.29	-.02	.13	.15	.34	.13	.17	.10
Stage 2	.29	.07	.14	.22	.32	.27	.19	.19
Stage 3	.32	.24	.07	.19	.34	.46	.19	.20
Stage 4	.32				.35			
All Data	.58	.25			.64	.49		

NOTE: Blanks indicate that ratings were not made for the judge-stage combinations in question. The numbers of cases are the same as in Table 12.1 (with a few very minor exceptions). Different judges saw different materials at the same stage (see Chap. 10).

Why should Liking correlate so uniformly well with the criteria? One possibility is that the ratings of Liking were tapping what Bingham (1939, pp. 221-223) called "valid halo" more than the predictions themselves. Let us listen to Bingham at some length on this topic. He begins with a general definition:

"The tendency for specific trait judgments to reflect in part the observer's general impression has commonly been regarded only as a troublesome constant error. The phenomenon, described by Wells in 1907 and christened 'the halo effect' by Thorndike in 1920, has been investigated from many angles.

"But even after observers have been cautioned against allowing general impression to masquerade in disguise of specific trait ratings, after they have been trained in drawing out and noting behavior indicative of the traits to be appraised, and after due allowance has been made for the positive correlations known to exist between favorable traits there remains a correlation between over-all evaluation of the person and specific trait ratings—a halo which cannot and should not be eliminated because it is inherent in the nature of personality, in the perceptual process, and in the very act of judgment."

Bingham goes on to say that the observer's perception of any particular trait varies with the configuration in which it is found; it is and should be rated as a characteristic of a particular person, who forms the background against which the trait is perceived as a figure. He describes a situation similar to our predictive study, in which judges had to appraise specific traits not in the abstract but as indications of personal suitability in specified situations. The observers therefore were under the necessity of comprehending a still wider ground or field which included the total personality pictured in relation to the duties or situations specified. This broad ground remaining constant, statistical evidence of halo in ratings made by skilled observers is to be expected and welcomed.

Valid halo, characteristic of correct ratings, should be sharply distinguished from the unwanted blur which marks a judgment as vague and undiscriminating, carelessly recorded when the observer's attention has been focused not on the trait in relation to its setting but on the ground alone.

To translate this argument into our own terms: Psychiatry is an occupation in which the general impression the practitioner makes on other people may be important for his effectiveness. Therefore, ratings which contain a large share of a global impression of "general goodness" or "likeability" may be the more valid for this halo content.

One thing to keep in mind about our Liking ratings is that they were made blind, *i.e.*, from the impressions derived from tests and interviews rather than from the direct impact of the person himself.\* Therefore, they may be quite different from ratings that the School's interviewers or testers might have made after seeing the man, possibly less influenced by factors such as appearance, superficial manner, and latent homosexual attractiveness or repulsiveness, all of which might have introduced more error (*i.e.*, variation unrelated to psychiatric effectiveness) into ratings of Liking (*cf.* Chap. 14). It may be this difference that accounts for the fact that the Liking ratings made by predictive judges in the Michigan project after direct contact with subjects failed to correlate appreciably with any criterion rating but were related instead to the Liking ratings of the criterion judges (Kelly and Fiske, 1951).

This last point suggests the possibility that both the predictions and the Liking ratings of the predictive judges in our study correlated well with the criterion because the latter was heavily loaded with likeability. It is true that the criterion raters' Liking for the residents correlates between .64 and .74 (Supervisors) and between .48 and .68 (Peers) with their evaluations of psychiatric competence (see Table 13.4). Perhaps all the apparent predictive validity is merely a dim reflection of the agreement between predictive and criterion judges that certain residents were "good guys" and others were "bad guys" or otherwise not likeable personally?

Fortunately, we have a direct measure of the agreement between "predictive Liking" and "criterion Liking" which can settle the issue for us. If the above hypothesis is correct, these two Likings should correlate with each other more highly than any predictor or predictive judge's Liking correlates with the criterion. But this is not the case (see Table 13.5). It is true only of Judge III (whose validities were generally among the lowest in the study) that his Liking ratings correlate at about the same level with the criterion raters' Liking as they do with Over-all Competence. For the other three raters, the validities of predictive ratings and of their personal Liking ratings are consistently *higher* than the correlation between their Liking and the criterion raters' Liking. Moreover, the predictive and criterion Likings are less closely correlated in the case of Peers' Evaluations—against which all validities were *higher*—than in the case of

\* So far as we know, Liking has never before been used as an assessment variable in this "blind" way.

Supervisors' Evaluations (see Appendix 13.1 for tables and full details).

We are left, then, with the proposition that all of these results make sense if we consider likeableness (in our special meaning) intrinsically related to aptitude for psychiatry. All of the predictive judges had personal friendships with psychiatrists and residents. It was to be expected that we would like those people we thought would be good psychiatrists; and of course we could see this happening as we were making our ratings.

Perhaps no further explanation is needed. But we are tempted to speculate a little more on the findings. We were psychologists, trying to decide which applicants would make good psychiatric residents; inevitably, in the background of these deliberations, were residues of our previous contacts with residents. We had worked with psychiatrists and knew something about the kind of person with whom working was a pleasure. Inevitably we *like* a colleague if working with him is gratifying. It should not be surprising, therefore, if a psychologist were to form an ideal image (perhaps not conscious or verbalized) of a good psychiatrist as someone who inspires liking as a working teammate. This conception, reflected in the Liking rating, might contain a good deal of validity as a predictor of actual performance.

Let us, then, look again at the thirty-two aspects of personality that were used as intervening variables in the predictive study and ask how we like people who possess such qualities in greater or less degree—especially, how likeable they are as professional colleagues (see Appendix 13.1 and Table 13.7). Surprisingly few of the variables were really neutral, at least for the predictive judges. Except at the extremes, perhaps a dozen of these qualities might make relatively little difference in our liking for a person, but the others would play a considerable role.

We like people because they like us—because they have a capacity to give love to others rather than being entirely self-centered or withdrawn (*cf.* Self-Objectivity, Need to Help) or always looking for admiration and attention (Status-Mindedness). (In addition, a status-oriented psychiatrist makes a particularly unwelcome teammate for members of other professions, who are likely to find themselves feeling patronized and downgraded.) Someone who is struggling to control his aggressive impulses and is likely to be surly, irritable, explosive, or sugary-sweet is difficult to like (Handling of Hostility,

Emotional Control). Genuineness, Sense of Humor, and capacity for good Social Adjustment with Co-workers are obviously appealing qualities. We like people with whom we can communicate, who have a capacity to listen and to be aware of our reactions to what they say as they are saying it; otherwise we bore each other or step on toes (Empathy). A person who has deep convictions, who is intensely absorbed in his work is usually attractive to others with generally similar values even if they do not share these particular interests and convictions (Internalized Interest in Work or in Psychiatry; Internalized Ethical Standards). The appeal of someone who is not personally static but on the move toward self-realization is similar (Capacity for Personality Growth).

When it comes to such matters as Psychological-Mindedness, Stereotypy of Thought, and Cultural Wealth, then we are probably getting out of the realm of the more universally important in human attraction and closer to values shared by the judges as psychologists. The same probably goes for Objectivity toward Dependent Figures, Consciousness of Social Injustice, and Freedom from Ethnocentric Prejudice.

These qualities have two features in common: (1) They are aspects of a person that have a lot to do with whether such people as one of our predictive judges will like him or not, and (2) they determine whether or not a man will be judged a good psychiatric resident. This pair of facts probably goes a long way toward explaining the high correlation between our ratings of Liking and the criterion variables.

Possibly they even contain a clue to the slight superiority of Liking over predictive ratings in forecasting psychiatric performance. Each judge had his own personal "formula" for a good psychiatrist, in terms of our personality variables, and, much less consciously, each probably had a slightly different formula for someone he would like—e.g., one judge might consider Self-Confidence a crucial asset for a good resident but find a candidate very likeable if he was creative, decent, and sensitive even though plagued by self-doubts. If then it turned out that Self-Confidence was less important than Capacity for Personality Growth, freedom from Stereotypy of Thought, Emotional Control, etc., his Liking rating would contain a better weighting of these qualities than his predictive rating. We suggest that if the differences in validities are not just statistical artifacts (sampling errors, or the like), they may be the result of this kind of process.

## FACTS OF PERSONAL HISTORY AS PREDICTORS

It has been found in various kinds of selection research during the past twenty-five years that facts of a person's biography may predict other facts about him as well as or better than tests or interviews. We never tried to develop a biographical inventory of the type used by the military services in selection studies (*e.g.*, see Rohles, 1953), in line with our general preference for projective techniques instead of objective self-administering instruments. Nevertheless, we looked into the possible validity of a few predictors of the type that are sometimes used in such inventories and related objective facts of an applicant's personal history.

**MEDICAL SCHOOL RECORD.**—At least a score of studies had shown that premedical-school grades correlated quite respectably with grades in medical school, especially first-year grades (see Stuit, *et al.*, 1949). The Menninger School of Psychiatry program, furthermore, was formally set up as a school with courses, practical and regular evaluations (though no grades in the usual sense). There was reason to suppose, therefore, that medical grades might forecast success in residency.

The main information we had about the medical-school record was the transcript of grades. Since it is well known that numerical and letter grades vary widely in their meaning from school to school, the measure used was the man's rank in his medical-school class. Out of over a hundred transcripts that were available, from only fifty-three was the rank determinable. This sample included a sprinkling of residents from all six experimental classes, however.

The correlation between this measure and over-all performance in the School was nil ( $-.04$ ). The absence of correlation can be taken to mean that even though psychiatry is a specialty of medicine, it requires a different brand of competence from the kind that gets a medical student good marks. At least one main consideration ought to be kept in mind in interpreting this result: The medical-school rank is largely a measure of academic competence, whereas the over-all rating as a psychiatrist is largely a professional performance rating. We have seen that academic performance in the School did not correlate significantly with Over-all Competence as a psychiatrist (*see Chap. 8*). It would be more reasonable to expect rank in medical school to predict grades on academic examinations during residency,

but both bits of information were available on too few residents to permit this hypothesis to be tested.

AGE.—The ages of the residents at the time they started training ranged from twenty-three to forty-four years. The large majority of the group were between twenty-four and thirty-five. The lower limit is, of course, set by the length of necessary educational preparation; it is difficult to complete medical school and internship before the age of twenty-three. Only 15 residents in the School were as young as this; most of them did about average work.

Off-hand, one might expect that since older residents had had more professional and life experience, they would perform better. The trend, however, was slightly in the opposite direction. Age at the time of entrance correlated negatively ( $-.16$ , not significantly different from zero) with Over-all Competence in the School. Moreover, out of the 230 Fellows, not one who was aged thirty-seven or over when he began training received a performance rating above average. Of the 24 men beginning in the School at the age of thirty-seven or older, 6 were rated in the average range, 7 were rated low-average, and 11 were considered inadequate.

The finding that the oldest students tend to do slightly more poorly has been commonly reported in medical education as well (Stuit *et al.*, 1949). Among medical students and psychiatric residents alike, the reason probably lies in the significance of a person's changing his occupation at a time of life when others have found themselves vocationally. Since this is an exceptional act, it implies exceptional and often unfavorable circumstances: perhaps such a man has found it difficult to commit himself to any profession, has not done well in his previous work, or has suffered some unsettling personal shock.

MARITAL STATUS.—Several expectations about the relation between marital status and competence as a resident in psychiatry seem justified: there should be no marked relationship, but one would expect unmarried residents, on the average, to be somewhat inferior to married ones, particularly in the older age ranges. All of these expectations were borne out by the data: *E.g.*, among 230 accepted residents, the 47 men who were single, divorced, or separated tended to be classified in the lower 40 per cent on Over-all Competence proportionately more often than the 183 who were married (51 per cent *vs.* 37 per cent; Chi-square significant at the .10 level). These trends

are not marked enough to be of much use in selecting residents, however.

A NOTE ON PREVIOUS TRAINING IN PSYCHIATRY.—Quite a few of the residents at the Menninger School of Psychiatry came after having had some psychiatric training elsewhere. Such men did not perform better than residents who had all of their training in Topeka, but it was easier to predict their level of competence. The validity of the average interview rating was .21 for those with no previous training, whereas for those with a year or more of training the validity correlation was .46. The difference in these correlations is of borderline significance ( $p = .10$ ), but it suggests that an interviewer can learn a good deal by questioning a man about his reaction to his first psychiatric work, his attitudes toward patients, psychiatric supervisors, etc.

#### SELF-ADMINISTERED INVENTORIES OF INTERESTS AND ATTITUDES

One type of test that was given a limited tryout in the Small-Sample Study was the structured, self-administering attitude inventory. Our lack of success with the inventories that were tried combined with preconceived preferences for less structured, "projective" tests and interviews resulted in our almost complete neglect of this type of instrument during the rest of the project. Only the Strong Vocational Interest Blank continued to be given to most subjects.

STRONG VOCATIONAL INTEREST BLANK.—The SVIB was one of the original battery of tests chosen for trial in the present research, though (like the Szondi test—see Appendix 13.2) it was almost never used in the actual job of selecting residents. Applicants were asked to fill out a Strong Blank (booklet form) at the time they were tested. By the time the test had been sent away for scoring and returned, the decision to accept or reject had usually been made. In addition, the psychologist-evaluators usually had little experience with the test, or faith in it. Nevertheless, because of its wide use and potentiality for vocational selection, it was retained on a research basis.

All SVIB's were scored on all available published keys and most of them on two additional, unpublished, keys. If we had any hopes at all for the Strong, they were fixed upon these two new keys. The first, a VA Clinical Psychologist key, was developed by the Michigan Project according to the standard procedure; it was based on re-

sponses to the Strong made by most clinical psychologists employed by the Veterans Administration in 1946. The other, called Psychiatrist A, was developed by Dr. Strong in a study of four medical specialties (Strong and Tucker, 1952) and incorporated the responses of about a thousand diplomates of the American Board of Psychiatry and Neurology.

Although in a preliminary study of the twenty-five residents rated highest and the twenty-five rated lowest in the first three classes, the VA Clinical Psychologist key gave a significant mean difference, it failed to correlate significantly with any of the major criterion ratings in our final Predictive Study of sixty-four subjects, as did Psychiatrist A. Curiously enough, the only keys to correlate significantly better than zero with any major criteria were the old Psychologist key (validities of .21 and .22 against both Peers' and Supervisors' Evaluations of competence in Diagnosis and Psychotherapy), which presumably represented the interests of "brass-instrument" experimentalists of the twenties, when it was constructed, and the Lawyer key (.22 and .23 with the two Psychotherapeutic Competence criteria). Even these two keys failed to yield a cutting-score that significantly separated satisfactory from unsatisfactory residents. (For details see Appendix 13.3.)

No doubt the ingenious quantitative methods devised by Strong might give better results than were achieved here if used under the most favorable circumstances: large numbers of subjects, good criteria, and keys made up for this express purpose. We believe, however, that he would have to change one important point of his procedure if the resulting key were to be used for selection. The method presently used is to take successfully established representatives of a profession, give them the test, and construct a way of scoring the responses that differentiates them most clearly from "men in general"—*i.e.*, people in all other occupations. In two ways, these are inappropriate types of data for this purpose; the test is to be used with younger men who have not yet settled into a vocation, and they are to take it under circumstances that arouse defensiveness and attempts to make a good impression, whereas the criterion groups Strong uses are under no such pressures. If it were possible to test thousands of interns as they were applying for all kinds of residencies, later following them up and keying their responses in accordance with the specialties in which they later established themselves, both of these objections might be overcome. Until then, it looks as if the

Strong Vocational Interest Blank will be of negligible value for selecting psychiatric residents.

**THE E AND F SCALES.**—The attitude questionnaires developed by Daniel Levinson and his associates in research on the authoritarian personality (Adorno *et al.*, 1950) are among the most widely used instruments of the objective type today. We tried out two of the tests developed by this group: the E Scale, designed to measure ethnocentric (*i.e.*, "racial") prejudice, and the F Scale, designed to measure fascist or proto-fascist attitudes. The two scales were presented in a self-administered questionnaire, with items intermingled: a 10-item E Scale and a 34-item F Scale. The scales were tried out with the Small Sample extremes; nine Highs and eight Lows completed the questionnaire.

Neither scale showed any promise whatever, being unable to differentiate these extreme groups. The general trend in both groups was very much in the direction of democratic attitudes and low ethnocentric prejudice; half of the members of each group received the *lowest possible* score on the E Scale. An item analysis failed to turn up more differentiating questions than would be expected from random variation alone (see Appendix 13.4 for further details).

The conclusion we drew from the above results was that a test consisting entirely of more or less obviously prejudiced, authoritarian, superstitious, reactionary, and other unprepossessing statements was too crude to differentiate even the extremes of a group, most members of which were highly educated and sophisticated. In addition, the generally liberal and progressive social atmosphere of the Menninger School of Psychiatry creates a climate of opinion in which a person with prejudiced or reactionary ideas would be very reluctant to admit them.

**PSYCHIATRIC ATTITUDE QUESTIONNAIRE.**—Dr. W. R. Morrow constructed a special inventory with several objectives in mind, one of which was to write items that would sound less obviously "wrong." Although it consisted primarily of statements the desirable answer to which was disagreement, nine of the fifty-two statements were positively oriented. Six areas were covered: somatic and constitutional bias, authoritarian submission and aggression, anti-intraceptive viewpoints, destructive overcriticality, stereotypy, and atomistic behaviorism. This test was administered to a total of sixty subjects, of whom

the eleven highest and eleven lowest were members of the Small Sample and thus represent extremes of an even larger group.

Despite the greater subtlety and pertinence of these items than those in the E and F scales, the scale and its six subtests all failed to show any validity. Scores made by the extreme subjects overlapped almost completely. Nor did enough of the individual items show enough promise to encourage further efforts to develop such an instrument for selection (see Appendix 13.4 for further discussion).

#### PREDICTORS FROM THE WECHSLER-BELLEVUE SCALE

David Wechsler's Bellevue Scale of Adult Intelligence, Form I, (1939) was the only test of general intellectual ability used in our project. It provides quantitative and reasonably objective scores which can be related to the various criteria, so we decided to try out a number of them in an exploratory spirit.

We had little hopes that any of the Wechsler-Bellevue measures would yield anything. Though it is the most widely used test of adult intelligence, it is considered to have little "top"—not enough difficult items to give an estimate of intelligence at the upper levels that is reliable enough to distinguish between, say, the "very superior" and the "absolutely brilliant." Then, too, we had looked into the various subtest scores and the IQ's of our Small Sample subjects and had found a great deal of overlap, with only slight trends toward better scores for the higher-rated residents. At one time, we planned to construct a special manual for the use of this test in selection, in which we were going to emphasize those qualitative features that make it such a valuable tool in the everyday diagnostic work of the clinical psychologist—particularly aspects of verbalization. We realized that intelligence was required for good psychiatric work but were impressed by several striking reversals of expectations in the Small Sample group. One of the best residents, for example, had a total IQ just under 120, which was one of the lowest obtained from an accepted candidate. Moreover, though he was clearly not as sophisticated an intellectualizer as many, this man appeared in no way slow or dull intellectually. Therefore, we did not even try out other, more high-powered intelligence tests.

It was easy to throw into the final correlation matrix the three intelligence quotients that Wechsler's test gives us: one based on the five *verbal* subtests, one based on the five relatively nonverbal *per-*

formance subtests, and the *total IQ*. The array of validities that emerged was quite astonishing (see Table 13.12, Appendix 13.5). Performance IQ did not do well, and Total IQ was in between, but Verbal IQ correlated significantly with every major criterion, as high as .47 with Diagnostic Competence (Supervisors' and Peers' alike). It predicted Over-all Competence just about as well as the School's tester could do with all the rest of the information in the standard battery in addition to this measure of verbal intelligence (.39 for IQ *vs.* .38 for tester—Supervisors' Evaluations; .36 for IQ *vs.* .47 for tester—Peers').

Such a promising predictor had to be cross-validated; if it held up, it would be the best solution yet found to the problem of selecting residents. Of course we had the same data for all residents in our sample, so it was no trick to compute correlations between Verbal IQ and Supervisors' Evaluations for the other subjects. The results were not as impressive as before. For the 128 subjects of Classes I through III, the correlation was .27, whereas Class IV ( $N = 27$ ) gave a validity of exactly .00. It should be remembered, however, that Class IV was somewhat freakish, anyway, in terms of most predictors; neither the School's tester nor the consensus of the interviewers could forecast the future performance of these residents better than chance. When we throw all subjects together, for a grand total of 219, the validity remains at .27, which is, of course, highly significant statistically, though not of marked usefulness practically. We should not turn up our noses at it, however; only one of the three testers for the Menninger School of Psychiatry, and only two of the nine interviewers, could do as well or better, and it is exactly the same as the average validity of psychological test ratings for the total group of subjects.\* Indeed, the test ratings correlate moderately (.40) with Verbal IQ and if the contribution of the latter to the testers' ratings is held constant by partial correlation, a barely significant residue of validity remains (.19).

To see how well Verbal IQ would have performed as a practical device for selecting residents, however, we must ask if it gives a good cutting-score, one that discriminates well between acceptable and

\* That intelligence and achievement in many fields are closely related has been amply documented by Terman (1954), who remarks: ". . . I am convinced that to achieve greatly in almost any field, the special talents have to be backed up by a lot of . . . the kind of general intelligence that requires ability to form many sharply defined concepts, to manipulate them, and to perceive subtle relationships between them. . . ."

unacceptable applicants (see Table 13.13, Appendix 13.5). Unfortunately for this purpose, Verbal IQ works best at the high end of the scale and is distinctly inferior to the tester's judgment in making the basic distinction that is important for operational use (*cf.* Tables 13.13 and 13.14). The best place to cut the distribution would be between Verbal IQ's of 120 and 119. Rejecting everyone with a Verbal IQ under 120 would have eliminated seven of the 55 Drop-outs (13 per cent of them) at a cost of only one Superior resident, six who were Adequate (3 per cent of that category), and one man of Borderline performance. Curiously enough, none of the men who did inadequate work as residents would have been excluded by this rule. Raising the cutting-score enough to eliminate any appreciable proportion of them would have sacrificed an inordinate number of Adequate and Superior men (see discussion in Appendix 13.5). Table 13.13 shows that the tester's rating of 4.0 or lower would also have picked up seven Drop-outs plus one Inadequate, but with the loss of *no* resident of Superior, Adequate, or even Borderline performance!

At the other extreme, if a candidate had a Verbal IQ of 137 or better, he was very likely to do well. Only 9 per cent of the Drop-outs scored so high, and only a single man each in the Borderline and Inadequate categories. For all practical purposes, therefore, if an applicant achieves a Verbal IQ of at least 137, he can be accepted without further testing or interviewing at little risk. Since only about 15 per cent of all applicants to the Menninger School of Psychiatry had IQ's this high, however, such a rule would not have been of great assistance in cutting down the labor of selection. Even at the high end of the scale, moreover, the testers' ratings worked better: Considering approximately the same proportion of highest-rated men, we find that *no* Drop-outs would have been included and only one Borderline and one Inadequate resident.

The cutting-score analysis shows, then, that the good correlation of Verbal IQ with Over-all Competence is due in great part to good discrimination in the upper ranges of proficiency, where it is little needed, and that the IQ cannot substitute for clinical judgment in making the decision whether to accept or reject. The practical significance of the cutting-scores is further diminished by three factors: (1) The results cited are based on the entire population and are therefore not cross-validated; one should not expect to do as well on a new group. (2) The Wechsler-Bellevue is widely available to ap-

plicants, and it would not be difficult for most physicians to memorize enough of the answers to score in the Very Superior bracket; the revised form (Wechsler Adult Intelligence Scale), which is replacing the Wechsler-Bellevue in general use, is equally accessible, and its equivalence for this purpose is unknown anyway. (3) If one tried such an instrument as the Miller Analogies Test, a measure of verbal intelligence that is *not* accessible for advance study, it would be necessary to repeat the entire study; it could not be assumed that any particular MAT score would work just as well as a Wechsler Verbal IQ of 137. Nevertheless, such an experiment seems to us well worth trying.

#### RORSCHACH SCORES AS PREDICTORS

Another class of possible predictors, thrown into the hopper with a spirit at once exploratory and skeptical, was a selected group of scores from the Rorschach test. In contradistinction to Rorschach predictors used in other parts of this project, these were the familiar sums and percentages that appear on the Rorschach psychogram as usually tallied, plus a few special combinations.

We were not surprised to find that most of the fourteen scores failed to give significant correlations with any of the criteria. A few of them looked fairly promising, with a scattering of significant validities, but the percentage of rare details stood out over all others—and, in fact, over every other predictor used in our entire study except Judge I's Final Stage ratings—in uniformly correlating significantly better than chance with every criterion except one (see Table 13.15 in Appendix 13.6). It correlates between  $-.36$  and  $-.44$  with each major criterion, and up to  $-.50$  with the minor criteria (at the 1 per cent level with ten of them, at the 5 per cent level with two others, using two-tailed tests).

These correlations mean that men of Classes V and VI who amassed many responses to tiny parts of the ink blots or to white spaces, bits of the edge, and similar unusual areas, tended to be poor at all aspects of psychiatric work. Clearly, if such strong relationships were to hold up throughout other classes, we should have a most valuable predictive device.

Fortunately, we had Rorschachs for almost all of our subjects, and we could correlate this promising-looking predictor with at least Over-all Competence for the four preceding classes; we did so at

once. The result was an almost complete reversal: for 116 subjects of Classes I through III, the correlation was .06; for 27 members of Class IV, it was .10. Notice that the sign of the correlation coefficient changed from negative to positive, even though it is insignificant in the earlier classes. This recalled the overlooked fact that one of the cues in the first Rorschach manual, based on the Small Sample, had been the percentage of rare details, but at that point it was supposed to be characteristic of *good* residents to give a high proportion of such responses. The cue had been dropped from the manual; it is easy to see why.

#### SOME ATTEMPTS AT ACTUARIAL PREDICTION

Actuarial predictive systems often employ groups of predictors, none of which has very high validity alone, and sometimes achieve useful results with them. We decided to try out some statistical combinations of the predictors that have been discussed in this chapter to see whether this approach offered any promise in the selection of residents.

Our first such effort used the Rorschach alone. The best-appearing Rorschach predictors were put into a multiple regression equation and cross-validated on the subjects studied earlier in the project. The five most likely-looking Rorschach scores gave a multiple correlation of .43 with Over-all Competence on Classes V and VI. When applied to 116 subjects from Classes I-III, however, the statistical prediction from the regression equation had a validity of —.04.

But perhaps it is not fair to the actuarial method to restrict it to scores from one test. Therefore, we tried a number of other combinations, drawing on the best predictors from the Wechsler-Belle-vue and Strong tests as well. A combination of the Verbal IQ, DR%, the Strong Lawyer key, and Strong Psychologist key yielded an  $R^*$  of .52 with Supervisors' Evaluations of Over-all Competence in Classes V and VI, and when the number of Good M responses was added, the  $R$  was .56. Both of these combinations failed to hold up under cross-validation; the prediction of the multiple regression equation correlated only .12 and .13, respectively, with Over-all Competence in the earlier classes.

After the fact, it is plain enough that a major reason for the failure of these formulas is their inclusion of the unstable Rorschach scores. Since such scores are not as objective as the others, we tried

\* Multiple correlation.

dropping them. The multiple correlation of Verbal IQ, Strong Psychologist and Lawyer was only .40 on the original group (Classes V-VI), but it held up better on cross-validation. With 155 cases from Classes I through IV, the predicted score correlated .24 with Over-all Competence, which is no better than IQ alone. As we saw above, the significant correlational validity of Verbal IQ lost luster when it turned out that no very good cutting-score could be obtained for use in actual selection; perhaps the addition of the Strong keys would give a usable cutting-score. It did improve the cutting-score validity of the IQ, but still not to a useful level: The cut-off was so low that although it was significantly better than chance on cross-validation, it was not significantly better than following the base-rate (*i.e.*, accepting everyone). (For details see Appendix 13.7.)

The best that our actuarial efforts could achieve, therefore, was nothing more than the Wechsler-Bellevue Verbal IQ, which retained at least a modest correlational validity in every class but one. Yet it was consistently outperformed by clinical ratings, which never had correlational validity under .12 and which did a better job of indicating which men should be rejected, which accepted.

Besides the obvious lesson of these statistical labors—that clinical judgment is a better method of combining data to predict such a complex type of competence as our criteria measure—there is a further moral. It is the same as the one taught by the marked shifts in the validities of the School's test and interview predictors from class to class, the same as the lesson learned from the disappointing performance of most manuals on cross-validation: Ignore variations in the sampling of subjects at your peril! In clinical psychology and psychiatry we are used to relying on intensive studies of a few subjects rather than on large groups drawn in a systematically random way from a specifiable population. For many purposes, this intensive method works, probably for two reasons. First, when you are studying a universal human characteristic, or one that is at least very widespread, sampling considerations are irrelevant. If you want to learn how many limbs a human being has, studying a few subjects picked up in almost any way will give a satisfactory answer. Second, the intensive study of single cases usually yields generalizations that are adequately protected by many contextual qualifications. The closer you stick to the rich concrete particulars of your data—as clinicians are wont to do—the less danger you will make rash general pronouncements.

It is when we turn to quantitative methods and trust in the complexity of our statistics, or their Fisherian descent, to protect us that the danger of ignoring sampling considerations becomes grave. Even staying within a single school and taking successive classes of residents, eliminating the most obviously "different" (by virtue of minority sex or ethnic classification), we found ourselves dealing with subtly, but unmistakably, different kinds of men. And in these different groups, functional relationships between quantitative variables (such as Rorschach scores and criterion measures) could emerge, disappear, or show up again in contrary and opposite forms. To discover consistently valid "objective" predictors would take much larger groups of residents, studied in this way over a long period of time, before the vicissitudes of sampling could become only minor ripples in the great group trends.

## A BACKWARD LOOK AT SOME PREDICTIVE ERRORS

WE HAD EXPECTED THAT THE GREATEST YIELD FROM OUR RESEARCHES would come not from positive findings but from studies of faulty predictions and the insights they might provide into the sources of error. It was not so easy as we had thought to look dispassionately for ways that assessments and forecasts had succeeded and failed because of limitations in the methods of doing so, but what we learned in the effort forms the contents of this chapter.

### SOME SOURCES OF ERRONEOUS ESTIMATIONS BY PSYCHIATRIC INTERVIEWERS

The results of the interviewing approach deserve special attention for several reasons: information was available from each interview about the psychiatrist's impressions of the applicant (on the brief form filled out by the interviewer; see Appendix 9.1). Many psychiatrists took part in the interviewing so that the sample of interviewers is not too small for generalization; and since interviewing as a method of selection is undoubtedly here to stay, we must make it into as good an instrument as possible.

There is less to be learned about the errors of the psychological

testers, largely because the sample is so small: two of the six psychologists did most of the testing, and one of these did about four times as much as the other so that impressions about the kinds of errors attributable to testing would be overweighted by the performance of that one psychological tester. Nevertheless, it seems that all of the testers reported more evidences of psychopathology than did interviewers, which may account for the testers' tendency to underestimate the potentialities of applicants.

We began our study of the interviewers by identifying the applicants about whom each had made the most erroneous predictions and then comparing the report of the interviews with Supervisors' Evaluations of how the men performed as residents (see Appendix 14.1 and Table 14.1). Taking men who were overestimated by one of the interviewers, for example, we found that they had in common (according to their supervisors' descriptions) superficial self-assurance which covered up many defects, especially impulsiveness, lack of judgment, and schizoid tendencies. Did these superficial impressions of self-assurance deceive this interviewer (Dr. D) and cause him to overlook the underlying deficiencies? We get some confirmation by looking also at the three men to whom he gave some of his highest ratings and who turned out to be Borderline. Two of these three also showed the same personality characteristics just described.

Qualities that were found to be responsible for the errors were often *not* missed by the predictors. *E.g.*: Interviewer D very much overestimated one applicant although noting, "He may barge into things too rapidly but will be aware of this difficulty." The quality of insensitivity described was very important in this man's failure as a resident; but the resident was *unaware* of the difficulty and unable to control it. Dr. D rightly credited another applicant with superior intelligence, breadth of cultural assets, and psychological-mindedness. Yet, in the course of his interview report he noted in three different contexts the man's inhibitedness. All the assets were perfectly clear in his later performance, but the interviewer did not see how pervasive and severe the inhibition was and how evident it would become as a chronic depression, seriously interfering with the resident's work.

Applying this method to the erroneous predictions of Interviewer C, we find that in his nine extremely overestimated cases, he tended to be overimpressed by two qualities especially: a man's size and physical appearance and the breadth of his fund of general in-

formation; the opposites of these qualities appear also in his three most underestimated cases. Like Dr. D, he saw a number of men's qualities essentially correctly, but a halo of general impression incorrectly raised or lowered the final rating. In a number of the reporting forms filled out by this interviewer there are signs that the contact was brief and hurried, indicating that the interview was superficial and missed many significant aspects, such as considerable psychopathology (in five of the overestimated applicants), or serious liabilities in the personality, such as overaggressiveness and insensitivity.

The interviewer who saw the largest number of applicants and who had the best general record, Dr. A, was not as much impressed by physical appearance as Dr. C, or by the amount of general knowledge the man had. His judgments seem to have been based largely on the application of principles that have proved valid; for example, he was impressed when a man could describe a case well in dynamic terms and showed strong interest in psychiatry. But he seems to have weighted some of these qualities too heavily; *e.g.*, two of the three men he badly underestimated expressed organic, rather than dynamic, interests to him and seemed not to be direct and genuine. When a man was frank and open in admitting his inadequacies and need for treatment, this interviewer was inclined to minimize evidences of psychopathology which later showed up (particularly depression). In two instances where this happened, as was true of Dr. C., the interview seems to have been brief and incomplete. Possibly this interviewer was especially attracted by certain positive oral character traits, overlooking the kinds of pathological potentialities they implied.

It seems to have been often true that if a quality of an applicant was powerful enough to be responsible for making or breaking a man's work performance, it was likely to have been noted by the predictors, although they not infrequently were unable to weigh correctly its importance. In some instances the halo created by a validly applicable quality, or an attractive but irrelevant quality, caused the interviewer to stop short of recognizing the man's potential pathological qualities or other shortcomings.

This first method can at best tell us the kinds of people who were hardest to predict in terms of the way they appeared later. It would be more useful to see if there were qualities in the applicants *as the interviewers saw them*, which were responsible for predictive failures. We therefore made a study of interviewers' descriptions of

applicants from their reporting forms.\* In Appendix 14, there is a full description of the method used for the four interviewers who saw enough applicants to make feasible its use, and Table 14.2 shows how it was applied to one interviewer, again Dr. D. The method enables us to compare the extent to which any quality observed and recorded by the interviewer was associated with high *predictive* ratings and the extent to which it was associated with high *criterion* ratings. If it distinguished between the predictive ratings but did not distinguish on the criterion ratings, we considered it as a likely source of error. Similarly, if it distinguished on the criterion ratings but failed to do so on the predictive ratings, then it might be something that the interviewer had noted but to which he had not paid enough attention. The comparison, then, is between the judges' conceptions of those who might become good *vs.* poor residents and his conceptions of those who *were* later good *vs.* poor residents. For this type of comparison it is sufficient to stay within the "private world" of the interviewer rather than ask what were the "real" qualities of the men he over- and underrated. (It would be interesting and valuable, of course, to know, also, what these real qualities were.)

It seems that Dr. D's conception of the good applicant laid special stress on the following qualities: empathy with others, insight into oneself, good speaking voice, liberal political and social views, poised appearance, warmth, a variety of recreational and social interests, and good work capacity.

How do these compare with the qualities Interviewer D saw in the men who really became good residents? The quality of warmth was apparently most misleading, for the differences between the criterion groups were much smaller than the differences on the predictive side (*cf.* Chap. 12). Voice quality was the second most important source of error with work capacity, appearance, empathy, and poise next in order. Thus, his misconceptions seem to have been that an applicant should demonstrate "very good" warmth (instead of "fair") and that voice quality, appearance, and reported work capacity were relevant, which they seem not to have been.

Interviewers A, B, and C show some of the same misconceptions. For Dr. A, the qualities best expressing his idea of a good prospect

\* Note that this analysis is limited to the interviewers' notes and that the latter were guided in large part by the interview report form. Thus, since the form provided space for descriptions of appearance, poise, voice, etc., it may have directed more attention to these features than they deserved.

were poise, warmth, empathy with others, and absence of psychopathology. In general, his conceptions were less mistaken than Dr. D's or the others; but again it was "warmth" that he most overemphasized.

Interviewer B's conception included poise, warmth, and a positive general appearance, and he tended to find more serious psychopathology in those he thought poor than did Dr. D. He was most misled by poise and next by warmth.

Interviewer C emphasized poise, positive appearance, and warmth. Like Interviewer A, his mistakes were relatively few, but again his idea of the importance of warmth was most out of line.

An additional point is true of all four interviewers: their conceptions of men who were thought to be good *vs.* those thought to be poor are quite similar to their descriptions of those who turned out to be really good *vs.* poor. Thus, predictions were correct fairly often despite misconceptions, and those about whom errors were made may not have been seen as too different from those who were predicted correctly. The brief notes written by the interviewers and our simple system of classifying them are too crude to pick up all the distinctions (including valid ones) that they might have made. Nevertheless, it is possible to see a consistent tendency for every interviewer to be *too much* influenced by the more superficial aspects of candidates: their apparent warmth, attractiveness, poise, voice quality, etc. One is reminded of Giedt's finding that the important element of an interview for predictive purposes seems to be *what the subject says*, and the purely visual aspects may give misleading impressions (Giedt, 1955). In this connection, see the evidence in Chap. 12 (e.g., Appendices 12.2 and 12.3) that the Predictive Study interviewer made better predictions from a study of typed interviews (with a little listening to recordings) than he did right after the interview itself when the visual impression of the applicant was still fresh in his mind.

#### THE JUDGEABILITY OF THE APPLICANT

All the predictors (interviewers, testers, and Predictive Study judges) tended to make errors on the same people. E.g., of the 208 applicants judged by the psychiatrists who interviewed for the School, 32 were overestimated; 18 of these were also overestimated by the School's testers.

It is well known (*cf.* Allport, 1937; Horn, 1943) that the qualities of certain people are harder to assess than others'. Some of the errors we make may have to do with "opaqueness," "variability," or whatever it is about the person that makes judgment difficult. But there is no way to tell, when several judges err on the same man, to what extent the trouble lies in the unpredictability of the subject and how much it is the fault of shared erroneous preconceptions about what makes a satisfactory resident.

Even when criterion judgments were made, after direct contact with residents' work, there were men about whom supervisors disagreed sharply (see Chaps. 8 and 11). In many of these cases, undoubtedly the observers were reporting correctly on different sides of complex or inconsistent residents.

Such a man was Dr. Capstan. It was predicted by the Admissions Committee's psychiatrists and psychologist, by Judge I and Judge II, that he would be in the poorest quarter of his class, and he turned up in the highest! (See retrospective studies of Judges II and I, below.) During his residence, Dr. Capstan applied for psychoanalytic training and was rejected—still another set of low predictive ratings. When, after completing training with the Menninger School of Psychiatry, he took a job at another institution, we heard he was doing very well there. But, a few years later, we also heard from an analyst who was supervising his psychotherapeutic work about his weakness in this function. The personal qualities that made it difficult for him to do well in intensive psychotherapy were easily picked up—too easily, for they caused most judges to overlook other excellent traits that enabled him to do outstanding institutional work in administration and therapeutic management. And the reports of his supervisors varied in accordance with differences in the kind of psychiatric work he was called upon to perform.

#### IMPROVING THE SKILL OF INTERVIEWERS

A sample has been given of the kind of thing that can be learned from analysis of each interviewer's errors. It raises the practical question of whether such information can be of use in improving interviewers' performances. We saw that the bases of error in prediction are usually fairly similar to the bases of success; it must be difficult, therefore, to correct so complex a judgmental process. Although many of the sources of errors we have uncovered look as if they would

be easy for a judge to correct, we know that this kind of judgment is not wholly rational (*cf.* Chap. 13, discussion of Liking), and one does not change one's feelings so easily. Certainly, however, it must be practically impossible for an interviewer to improve without making a study of his errors. Under the usual circumstances, he gets feedback too late and too unsystematically. Practice without immediate reward or punishment (knowledge of results) is likely to stamp in errors along with good ideas: probably an interviewer develops a style of his own, but he is not likely to improve markedly.

By and large, the same view must apply to the question of the ability of the best of the interviewers to teach others how they achieve their superiority, for the best ones were not distinctly better than the poor ones, and the skill is so complex that it must be hard to put into words or communicate to others by any means. Yet such subtle qualities as empathic ability and judging facial expression are known to improve with practice. Taft (1944, p. 12) concludes, "There is good reason to believe that ability to judge others can be improved by specific training in judging and repeated specific practice, except where the person already has good ability to judge others." Reviewing one's past predictive reports on applicants after criterion information is available should be well worth the effort.

#### SOME RETROSPECTIVE STUDIES

If individual judges are to improve their skill in predicting, therefore, they should have the opportunity to study their own past performance—their errors and successes, after criterion data are in. To some extent it has been possible to try this step. Four interviewers and two of the Predictive Study judges were told which cases they overestimated, underestimated, and correctly predicted, and they looked back over their notes on these cases. It is some measure of how difficult it is to make such a retrospective study that one of the four psychiatrists concluded mainly that the criterion judgments were less correct than his own predictions, and three others delayed returning the forms for many months. Whether this kind of retrospective study leads to real improvement or not we do not know; certainly it produces a subjective sense of having learned something. Here follow three statements, one by an interviewer and two by Predictive Study judges on what they feel they learned from going over their errors.

The first are the comments of the psychiatric interviewer, Dr. D:

It would appear that I tend to rate a bit high. For example, I rated one man 7, which means average, but I said of him, "On the whole I didn't like him very well. I wonder if he will have tolerance for illness and if he can have much respect for his dependent patients. He did poorly at describing his parents and wife and did not demonstrate much feeling or psychological-mindedness." Looking at these statements it's hard to justify a rating of 7. It probably should have been 6 at the very most but apparently at that time I was taking 7 as a kind of an average mediocre rating. Those men whom I tended to overestimate, for the most part, I rated highly on points six, seven, and eight of the interview report form. I was impressed by their cultural interests, ability to express definite social and political awareness and views and seeming psychological-mindedness in describing the important people about them. Although I recognized in these people such things as inhibitedness, cautiousness, rigidity, passivity, immaturity, etc., I didn't weight them very heavily. I think in looking over these reports again I tended to minimize definite historical evidence of poor family background or even bad family history, neurotic parents, alcoholic father, etc. I took the man strictly on his own merits as I saw him and was most impressed by those people who had a wide cultural background, a breadth of social and political views, and intellectual ability.

I can't make much out of the three men that I underestimated. One of them I thought was breezy, crude, swore too much and was insensitive. He also could not tell me why he wanted to go into psychiatry. The other two men I rated "mediocre-six" but voted to reject them because I felt at that time that the number of applicants available allowed us to reject mediocre men. I didn't rate either of them down very severely. In both cases I stated that the applicant had some outstanding potentialities but I apparently was poorly impressed by what I called overcontrolled, overpoised or overintellectualizing traits with underlying aggressiveness.

In Dr. Himmel's case I got a good history of his being a severe misfit and an introvert in childhood. I still don't understand why he turned out as well as he did.\*

One of the Predictive Study judges wrote out his impressions of the causes of his over- and underestimations after studying the cases involved:

I note first the larger number of underestimations (9) than overestimations (5).† It reminds me that the psychological test predic-

\* Actually, his supervisors also had a wide range of opinion on his competence.  
† The figures given derive from a simple definition of misestimation: a difference of two quartiles or more between predictions at the Final Stage and the criterion (Supervisors' Evaluations of Over-all Competence) for Classes V and VI.

tions were similarly askew, and that it is a common complaint about the psychological tests in general that evidences of pathology are easier to see than the patient's strengths. Yet the battery I depended upon included a recorded two-hour interview which I went through rather carefully so that I am willing to consider that the surplus of underestimations has to do with a proneness on my part to underestimate people.

Like others whom we have asked to go over their errors, I find it difficult to discern their probable causes. I'm even prepared now to repeat the heresy that for some residents the School may be seeing with less perspicacity than I did. Most of the inferences in the survey of my errors that follows below are based upon what the supervisors say in their evaluations of the residents in comparison with what I emphasized in my predictive study.

In my five overestimations I strongly suspect that in all these cases very superior intelligence and knowledge overimpressed me as it often does so that other defects were minimized. In two cases the same might be said for my admiration of their verbal facility. According to the supervisors, all of the five residents had poor relationships with patients, four of these were poor because of inactivity, passivity or apathy, and distance, and the other one varied between being overactive and reluctant to be firm. In going over the cases I was struck by the difficulty of being objective about it. My attention may have been drawn to these aspects of the supervisors' lengthy descriptions by my awareness that at the time I made these predictions these same criticisms might have been applied to my own work.

The nine underestimated residents are rather more difficult to generalize about. I tried in my very general summary of what the supervisors emphasized to avoid distortion, particularly by concentrating on the qualities that were most frequently mentioned. The supervisors say of these men that (rather than being distant) they show considerable warmth, are liked by their patients, are very interested and enthusiastic about their work, they learn readily from their supervisors, and their most common fault is overidentifying themselves with patients. They are often described as making great improvement or having been initially underestimated (for example, cases No. 1 and No. 19). So far, though, these qualities stressed by the supervisors do not strike me as much different from those that are common for almost all Highs. I noted in my predictive statement and the supervisors agree that these are not men who are outstanding for their knowledge and intelligence. I may have reacted too negatively to that lack and not positively enough to their evident warmth to patients, enthusiasm for psychiatry, and ability to profit from the offerings of the supervisors.

My predictive foibles are especially evident in regard to Dr. Capstan whom I judged to be in the lowest quartile and who ended up in the top quartile of the residents. In my notes after reading the

interview and the earlier two stages, I noted that he was lacking in his ability to express himself although he seemed to be aware of this deficiency. I noted that he was not especially bright and seemed to be very disorganized in his work pattern. (I was wrong in this. He was extremely conscientious and hard working and supervisors described him as the work-horse of the ward.) After reading the psychological tests, I emphasized further his thinking difficulty and especially how easily his immediate memory and concentration were impaired by anxiety. His difficulties in expressing himself and his being not too bright were confirmed over and over again by the supervisors' observations. Of the various work functions I thought he'd be poorest in psychotherapy and after reading over his supervisors' estimates, this appears still to be correct. Unfortunately, in making my predictions I too generally assumed that if a man couldn't be a good psychotherapist, he couldn't be a good psychiatrist which obviously isn't true of this man. It was emphasized by the supervisors that he was very eager to do the right thing although his approach to the patient was not based on very great understanding. It was commented that he conducted his treatment in psychotherapy in a haphazard way but that his real assets were his eagerness to do this kind of work, his sincerity, reliability and conscientiousness, the fact that he can be a father figure to his patients, that he can maintain a good balance between firmness and sensitivity. While I noted in my predictions of the final stage that all recommenders agree in liking him and some were very enthusiastic about him, especially about his conscientiousness and interest in patients, I did not allow this to raise my rating from the third to the final stage as much as it deserved.

Quite late in Dr. Capstan's training, one supervisor cast up the balance in a way that shows how the defects that I noted were more than compensated for by assets. "He is an individual who can accomplish almost any kind of work demand without seeming pressed or without seeking out special credit for it. He really tries to relate himself to the needs of the patient at the expense of his own comfort and time. He has an enveloping kind of interest and warmth to patients but his greatest lack is in imaginative flexibility. He has mechanical concepts, he sees things best in concrete mechanical terms, feels more comfortable working with EEG data rather than the nuances of a transitory relationship. Although he has become to a certain extent aware of the nature of this difficulty, he has been able to effect but little change in it and his particular deficiencies at present are really not different from what they were at the time he came to this section. He makes up for his lack of flexibility by a kind of sincerity, of interest, an actual devotion to the problems of his patients. He gives occasionally a surface impression of being rather devoid of feelings but this is a mask behind which exists real sensitivity. With his patients as well as in his professional supervisory contacts he is at times himself unable to penetrate this mask. In this

way one is at times surprised at the amount of feeling that has been invested in the situation of which he at the time gave no indication."

It made the trend of my errors more obvious to me when I found them again on going over briefly the next resident who I predicted would be in the bottom quartile and who turned up in the top quartile. The supervisors said of him that he was young, sincere, conscientious, quiet, not sure of himself, lacking in experience, inhibited, but that he was initially underestimated. Although he frequently overidentified himself with patients, he was curious, he wished to help the patients, they liked him, and his improvement was gratifying to supervisors. True to form, in my predictions I emphasized the presence of some schizoid disorganization of thought with a paranoid flavor, that he was not especially bright, that he was very conventional and conservative in his outlook. I noted about him, though I did not weight it enough, that although his letters of reference were brief and unenthusiastic about him, they said that his assets were his conscientiousness and his interest in the patients assigned to him. Yet, of the several work functions, I thought he would, like Dr. Capstan, do poorest in psychotherapy and best in management and administration. Also in my initial impression and before I'd seen any of the test material, when I had only spoken with him a few moments, I was unimpressed on account of his apparent immaturity, conventionality and lack of sophistication about psychiatry and psychological tests.

Another Predictive Study judge wrote a similar retrospective study, which follows:

To begin with, I find I have to protest that not all of the disagreement between my predictions and the criteria is error on my part. I firmly believe that I evaluated one "overestimated" applicant pretty accurately, seeing his faults, but predicting that his virtues would shine through and he would grow—as he did. His ratings just did not quite catch up with him.\* About the same thing is true of one of my "underestimations," too; I rated him as barely passable, and on the whole, that would seem to be the consensus of his supervisors at the end, but he got so many would-be-noncommittal ratings in the low average range that the quantitative criterion doesn't give an accurate picture. There are three other examples in which I think the criterion probably ought to move a little in my direction, but such adjustments would still leave a good deal of error to be accounted for.

Secondly, I am impressed by the fact that six of the ten most overestimated residents were in the School only one to one-and-one-half years; this was true of *none* of my ten most underestimated.

\* This was Dr. Abbott; see Chap. 11.

Consequently, the basis of the criterion ratings is scantier than in other cases, and they had to reflect a man's initial difficulties in adjusting to psychiatry as the criterion ratings based on the full three years did not. We know that many of our men had difficulties during the problematical first year, and yet finished strong; late-bloomers are considerably more numerous in our population than fizzle-outs. It may well be, then, that at least a few of these men would have performed much closer to expectations if they had had a chance to finish. Several had to go into the Army; a few transferred to other programs for personal (usually financial) reasons, and one was really dismissed. But I simply missed the boat on him, and I know it.

Another possibly face-saving consideration: the incidence of contamination of one sort or another seems rather high in my 20 worst-predicted subjects, particularly the overestimated ones. Most often, the contamination was spurious: I'd be reminded of someone *other* than the subject himself, and my estimate of him would accordingly be pulled slightly away from what it would have been otherwise. Another kind of contamination error was recognizing a man about whom I had formed a false impression. I had had a social contact with one of these residents, for example, causing me to like him and think he was probably quite a good psychiatrist, and in going over his tests, I recognized him. Even though I tried not to be influenced by my contaminating contact, I probably was; his criterion ratings were quite low. Yet he is one of the ones who was drafted early in his second year, just when he was beginning to find himself.

About one rather badly overestimated subject, I wrote a devastating final sketch, in which I predicted all sorts of dire possibilities (including suicide); I described pretty accurately his many weaknesses and his few strengths but didn't reflect all of this adequately in the numerical rating I gave! I had started with a fairly favorable impression and had simply not lowered my ratings enough in the successive stages of analysis. In a number of instances, I started off like this by being unwilling to deviate very far from a prediction of average performance and didn't modify my ratings enough, as evidence accumulated. (In other cases, however, I went too far, so this may not be an important general principle. But it does raise the general question of artifacts of the procedure we went through: analyzing the material in each folder in a fixed order, which may not have been at all an optimal one.)

When I examine the changes in my ratings through the four or five stages on each subject, I find that there are about a half dozen on whom I was never near the criterion. At other times, I fluctuated as I got various kinds of data; the stage at which I was furthest off was the one where I got the TAT. In four out of the ten underestimations, I lowered my ratings considerably after going over the TAT, then usually raised them on getting further data, but not enough. And in two of the overestimations, I was furthest off at the

TAT stage. (I was *closest* to the mark several other times after seeing the TAT, however.) The most misleading source of data for my overestimations seems to have been the Self-Interpretation. It was the point at which I raised three of the overestimated men the most. Apparently the ability to give a frank, sensitive, and insightful analysis of one's own TAT stories isn't *necessarily* a sure sign of a first-rate psychiatrist—though it often is.

In general, the TAT usually exposed a man's weakest points, without giving compensatory signs of his strong ones—or at least without giving them in a way that I could read, even after the fact. To some extent the same thing is true of the Rorschach: I lowered ratings too much after the Rorschach more often than I raised them too much. We know that latent conflicts show up in these tests much more plainly than the compensating strengths; this experience impressed on me the necessity to be very cautious in assuming that such *potential* liabilities are *actual* if they are not seen operating in much more direct fashion.

Many of the men I underestimated had marked personality defects but high I.Q.'s, and they were able to do so well in the relatively intellectual aspects of psychiatry that they were rated much better than I expected. If a resident could grasp the intellectual content of his courses quickly, discuss the reading intelligently in literature study-groups, turn out literate write-ups with a good grasp of the diagnostic points, and could talk sensibly and informedly when participating in a psychotherapy group control, he could go a long way on brains and relatively little else.

For example, one resident seemed to me very nearly schizophrenic, and he was described frankly as such by one supervisor, who added, however, that because of his intellectual brilliance the man might do all right in psychiatry; he said he knew of several such colleagues in the field. Actually, almost all supervisors saw that this man was very sick, and he came within a hair's breadth of being asked to leave but was kept because people were impressed by his intellect and thought he would develop. Well, he did, from a level of almost total incompetence to a barely passing level when he left the School; very recently I have learned that he has a responsible position in a hospital, as well as a flourishing private practice, and by all external criteria seems to be doing fine. He, incidentally, got treatment after his first year of residency. The other thing that I missed about him was his tremendous energy and capacity for work—stemming partly, I think, from a rugged physique. I could not take this into account, of course, in a blind analysis.

This matter of constitutional energy and vigor seems to help account for some of my underestimations. Four of the ten were unusually big, muscular fellows with iron constitutions, and a fifth was not far behind them; only one of the overestimated fits this description at all, and ironically enough, he is one whose name was acci-

dentially left on one of his letters of recommendation, and I knew him by sight, so that I could take this asset into account! Actually, in his case I made a note to the effect that his physique was probably one of the assets that enabled him to function at all; it may have contributed to my overestimation. But when I didn't have the picture of the man to go on, I failed to predict his great energy and capacity for work—all underestimated men of this type were described as "real work-horses." Three of them form a kind of type: warm, earnest, hard-working, willing to stay late or come in anytime for their patients' sake, likeable—but at the same time called "not brilliant," mechanical and concrete in their thinking, not very psychologically-minded, and not good at subtle nuances or deep interpretations. In short they presented themselves as good decent human beings on whom one could rely, and made fine institutional psychiatrists but were notably poorer at intensive psychotherapy.

In their tests, these men all showed marked evidences of disturbed emotional adjustment (which was *not* overtly manifested), having to some degree schizoid and paranoid features and a good deal of latent homosexuality. Moreover, they gave signs of marked anti-psychological defenses, seemed poorly acquainted with their own feelings, and were not introspective. This probably contributed to their good external adjustment: being unaware of their conflicts, they were able to maintain an impressive front and get a lot done, getting by well as long as they didn't have to use their inner resources. Thus, they did relatively well in management and administration and poorly in psychotherapy, while my own over-all rating was too heavily weighted with psychotherapy.

Being an intellectual rather than a football player myself, I had a rather negative reaction to these extroverted, uninsightful fellows who often expressed more interest in sports than in cultural matters. Correspondingly, there were a couple of ectomorphs whose cultural wealth helped me overestimate them. I may also have been too favorably impressed by signs of real nurturance and empathy.

Several times, I laid too much emphasis on poor self-confidence. Though I think that I was usually not too far off in my estimates of it, self-confidence turned out not to be as essential as I had thought it was for psychiatrists. In a way, this is another instance of insufficient allowance for a good front covering inner weaknesses.\*

Another of my blind spots was for the positive uses of authoritarianism. Reaction formations and moralistic control worked better for a number of subjects than I thought they could. If a man had a strongly authoritarian character structure, I was likely to rate him low. I can recall several whom I predicted quite accurately on this basis, but there is such a high concentration of authoritarian char-

In a personal communication, Roy Schafer adds that "as a base for empathy with patients, some significant insecurity can be very helpful—and for insight, too."

acters among the men underestimated (nine out of ten are more or less of this type) that it was probably a source of error. A couple of times, I overestimated residents while seeing that they did have definite authoritarian elements in their make-up; the compensating features I saw didn't always operate as well as predicted. And one of them is a clear instance of my leaning over backward to compensate for my dislike of the man.

In two overestimated applicants, the main trouble seems to have been rather simply that the men were crippled by overobsessiveness, which I did not assess properly. In terms of the usual test signs, neither man was extreme for this population. In both cases I saw the obsessiveness but thought that it was not so badly out of control as it proved to be. Yet, I would have thought that the battery of procedures we used would give an excellent measure of just this aspect of personality and psychopathology. Why didn't it? In one case, I think the simultaneous presence of considerable inhibition held down the usual test indicators of overobsessiveness; in the other, I apparently erred in thinking that his huge Rorschach record, characteristically padded with responses to small and rare details, showed in other ways that his obsessive defenses had not decompensated and were well-controlled.

My general conception of what kind of person makes a good psychiatrist contained one serious error: I thought that free expression of affect was desirable, and overcontrol undesirable. Actually, as we have pointed out in several places in this book, the opposite tended to be true. A particularly good sort of pattern is marked sensitivity (which may even appear to be vulnerability) on the receptive side, combined with inhibition or overcontrol of emotional expression. This is found in half of the applicants I underestimated and in only one or two of the overestimated ones. Along with this failing might be mentioned the difficulty I had in assessing emotional control. In several cases I seem to have mistaken test signs of *impulsiveness* for *warmth*, or vice versa.

In several folders my notes record my having considered attributing a quality to a man on the basis of usual diagnostic testing principles, but refraining because I had seen similar cases in which the indicator had failed. In all such instances that I noted, I would have been better off to have stuck with the conventional interpretation, remembering that an exception does not necessarily invalidate an actuarial rule.

It strikes me in going over these, my worst predictions, how many of these subjects had had some prior knowledge of projective tests. If a man has read a book on the Rorschach, or has had lectures on the TAT; if he is already familiar with the plates and has some ideas about how he "ought" to respond (no matter how distorted they are), it introduces a big unknown. I was much more confident before we began that I could allow for this, that the real personality

could not be concealed, etc.—much more than I am now. The man I overestimated most badly knew a fair amount about projective techniques, enough so that he looked as if he had “counterbalancing strength,” along with the unconcealable evidences of deep disturbance. It was only when I got to the interview that I realized what kind of fellow I was dealing with, and at that point I was unfortunately contaminated and couldn’t make predictive ratings.

On the whole, I am impressed by the *qualitative* agreement between my sketches of the most poorly predicted subjects and the supervisors’ descriptions of them. Almost always, the most important assets and liabilities cited by the supervisors showed up in the tests, interview, or other materials. The real problem was almost always to cast up a balance—to figure out which things would endure, which would be outgrown; to gauge which trends would remain latent and which would become manifest; to estimate the efficacy of this defense against that conflict, or one asset as a counterbalance to the other liability. Here is where art, experience, and luck have to take over, for I think that we learn little from our texts or teachers in psychology about such matters. Many of the men whom I rated quite differently than the supervisors did were unusually complex mixtures of promising and unpromising features, and usually their work was colored in both ways—only not to the shade I had mixed up for them. In large part, this amounted to an inability to judge accurately a man’s capacity for personality growth. I am now convinced that to do this well requires careful assessment of all types of data we had.

#### CONCLUSIONS

Looking back is proverbially easier than looking forward, yet it may also make the forward glance clearer in the future. When retrospective vision was applied to the predictions made in Designs I and II (the Predictive Study), many limitations of both research designs began to be revealed. A great variety of reasons may be responsible, when predictive and criterion statements do not agree. (In Appendix 14.2, we have listed and discussed as many of them as possible.) The most important of them can be covered under three headings, however: (1) difficulties in the predictive process and fallibilities of the judge; (2) difficulties arising from the person whose future is to be predicted; and (3) weaknesses in the criteria being predicted.

1. One is impressed, in reading over the retrospective studies, by how often the judges complain of having assessed qualities correctly although weighting them wrongly in predicting future competence. There was some erroneous estimation of isolated qualities (*e.g.*, over-

obsessiveness for one of the Predictive Study judges), but more often the judges seem to have sized up a man correctly in terms of qualities that were not considered by the supervisors to be crucial. This is particularly plain in the analyses of interviewers' report forms, where it seemed that predictions were much influenced by superficial attractiveness (poise, voice, appearance, and especially warmth) while these aspects of personality were not related to performance. The psychologists, on the other hand, were too much influenced by signs of pathological trends in well-functioning people. Thus, in general, judges seem often to have been thrown off the track by only partly conscious preconceptions.

Several other difficulties inherent in our experimental procedure also emerged: faulty use of rating scales (Dr. D); failure to sufficiently modify ratings after forming an impression at early stages of analysis even though important information became available later; incorrect weighting of different work-functions in evaluating Over-all Competence; reluctance to reject enough applicants (see also Chap. 9); etc. (See Appendix 14.2.) Moreover, an element of soothsaying had to accompany all predictions. The judge was forced by the nature of the experiment to rely upon the tests and interviews to give a picture of the applicant in the present and past, and to predict his future behavior in situations that could be foreseen in only the roughest outline.

2. But the person himself was not only a sometimes inscrutable object of study, he was constantly changing. The impact and occurrence of psychiatric treatment, marriage, children, divorce, and other major life events could not be known in advance. Furthermore, the age period involved is one of considerable change in most residents, as we have seen in Chap. 5, particularly in certain residents such as Dr. Abbott (Chap. 11). In the retrospective studies, judges noted the difficulty of evaluating capacity for change and the *rate* at which correctly predicted types of change would take place.

3. Finally, discrepancies between predictive and criterion ratings were sometimes the fault of the criterion, at least in part. We relied mainly on the judgments of supervisors during training and have reason to think they constitute a good criterion. But they are not perfect measures of competence, and they varied from the time of our Small Sample studies, in which hypotheses were formed and manuals written, to the time when the final criterion measures were obtained to check the predictions. For some of the measures, reli-

bility was excellent but not for all; a fair amount of disagreement between supervisors remained, and the judgments on some residents shifted from time to time. There is reason to think that supervisors' agreement on some residents was a shared halo of common misconception. Factor analysis (see Appendices 8.12 and 14.2) indicated that supervisors were making highly global judgments of competence—as the Predictive Study judges were too, despite efforts to make ratings in concrete and specific terms.

In the presence of these and other hindrances, it may seem a wonder that any correct predictions were made at all. It should be kept in mind that most of these considerations bear on the effort to predict fine degrees of competence, which was not very successful, and have less influence on gross judgments of competence or incompetence, which were for the most part very successful indeed.

P A R T I I I

THE PERSONALITY  
OF THE  
PSYCHIATRIST



A SURVEY OF EXPERT OPINION ON  
PERSONALITY REQUISITES  
FOR PSYCHIATRY

ALTHOUGH THERE HAD BEEN NO ORGANIZED RESEARCH OF COMPARABLE size on the selection of psychiatrists or similar professional personnel before our project, the field was far from wholly untrod. The selection of psychiatric residents had been going on for many years: those with the jobs of selecting had opinions about it based on a good deal of experience. It was an obvious first step to examine these opinions. At the time systematic work on the present project began, a relevant summary of opinion had already been published in the *Bulletin of the Menninger Clinic*, in which the faculty of the new School of Psychiatry pooled their experience and insights in a statement about requisite qualities of personality in a prospective psychiatric resident. But, so far as we know, this was the only publication on that topic up to that date.

Several years later, Dr. Karl Menninger made a survey of most of the large psychiatric training centers in the country, inquiring about their methods of selection.\* The emphasis of this inquiry was mainly on procedures, but one question was germane to our present discussion: "Upon what qualities or evidences do you rely most heavily for your decision?" In answer to this question ten of the

\* Other results of this survey are described in Chap. 1.

seventeen respondents mentioned specific qualities of personality or character they sought in candidates for residencies.

In Table I-15.1 (p. 267) we have summarized opinions revealed by these two sources. Later, more material setting forth the opinions of other workers in the field became available, and we include them also in the table. These additional sources are: a report of the Committee on Medical Education of the Group for the Advancement of Psychiatry (Greenhill *et al.*, 1955); and a statement prepared by Dr. K. A. Menninger for the Conference on Graduate Psychiatric Education of the American Psychiatric Association (1952).

In the same table, paralleling the list of qualities to be sought, or avoided, in candidates for training, we summarize the qualities specified as characteristic of good psychiatrists in three recent writings: a book by Redlich and collaborators published in 1953; an article by Ruesch comparing different types of psychiatrists, also in 1953; and an unpublished manuscript by Henry A. Murray, "Criteria for Selection of Psychiatrists to Work in Veterans Hospitals." (See also Appendix 15.2.)

Taking the table as a whole, one is struck by the number and variety of features of personality named. Yet Murray is the only contributor to the table who names as many as a dozen qualities. Thus, the superficial impression may be one of a lack of agreement. In the column dealing with applicants, only six out of the more than forty entries are named by as many as three contributors, and the only feature named by more than three is the most obvious qualification: a good record in medical school. In the column dealing with trained psychiatrists, no single entry is identified with more than two sources except "good technical training." But this apparent scattering of conceptions is due to the way they have been tabulated. It would have been possible to combine categories by making assumptions about what was intended, but it seemed desirable to hold such assumptions to a minimum. Terms such as intuition, empathy, understanding, and sensitivity do tend to be used more or less interchangeably by many psychiatrists (*cf.* Murray: "Insight: sensitivity to the subtler dynamics of human behavior, refined intuition similar to that of the expert novelist, accurate diagnosis through empathy"; all given as a single criterion), but they are used differentially by some and so were tallied separately here. On the whole, however, there is a great deal of internal consistency in the composite picture of the promising applicant.

There was one area of clear disagreement. On the very concrete point of performance in medical school, some psychiatric teachers said they preferred men who were in the top half or third of their classes scholastically. Bellinger wrote, however, "I prefer those in the last part of the first third or the second third of the class. Those with the highest grades are likely to be bookworms who are introverted, are not good teachers, and do not make good psychiatrists." Ewalt expressed a somewhat similar point of view, even though he said that "we do not like to take candidates who were not within the upper third of their group scholastically. We also try to pick people who have had a variety of experience—that is, those who have been interested in cultural, political or business activities in addition to their medical school work. We feel that this gives us a person who has interests outside of the scientific aspects of his work and who has some appreciation of the realities of the other people around him; it also avoids taking the stupid 'grinds' who can sometimes make fairly high marks in school." Gildea commented: "We have found that aptitude for psychotherapy does not correlate very highly with achievement in medical school."

On only two other points is there any evidence of possible disagreement. Bellinger states clearly that he wants extroverts, not introverts; though no one else uses that terminology, some of the desired personality types described by others would seem to imply an opposite emphasis. Compare, for example, Knight's stress on richness of cultural background and psychological-mindedness: the capacity to think about the self and others in terms of feelings, motivations, conflicts, etc. Finally, there may be some disagreement on the issue of freedom from symptoms. Grinker reported that there was a good deal of variance in his own group on this score: "One member of the selection committee will take the sickest and the most sensitive. Another will want the most stable and perhaps the most compulsive." Note also Redlich's remark that the psychiatrist "doesn't have to be a paragon of mental health."

A second general impression that the table gives is the high degree of agreement between prescriptions for the good candidate and those for the good psychiatrists.\* True, there are 18 entries in the

\* The extent of agreement would appear higher if the decision had not been made to tabulate the qualities mentioned by Menninger on only one side of the table or the other. Since his article was, in large part, a description of the kind of man "the psychiatrist" should be, everything listed on the left attributed to him should be understood to be on the right also.

right hand column of Table I-15.1 that do not appear in the left, but for the most part they seem to be individual variations of themes accepted very generally in the total group of sources (*e.g.*, common sense), points undoubtedly taken for granted (*e.g.*, interest in psychiatric subject-matter), or certain matters of relatively peripheral significance (*e.g.*, speed of work, perseverance) taken into account by Murray's list, which clearly strives after comprehensiveness, and not taken into account by answers to Menninger's question, "Upon what qualities or evidences do you rely *most heavily* for your decision?" [Italics ours.]

Nevertheless, a few qualities of the trained psychiatrist mentioned here are undoubtedly a function of his education. Consider the group of inner-personal capacities having to do with control of affective reactions in the work situation: Redlich says that the psychiatrist should have "considerable resources of self-control," adding that "scientific detachment is one of the doctor's chief tools"; Murray lists "Objectivity: separation of facts and sentiments" and "Personal detachment: can the man remain emotionally unimplicated?" Menninger makes a similar point at somewhat greater length: "He must have such supernormal self-possession that even the most bitter personal attacks of a patient are felt neither as insult nor injury, but rather as evidences of a kind of perverted trust." It is probably not accidental that no such desideratum was listed for the potential psychiatric resident. Doubtless it would be fine for him to have already such self-possession, objectivity, and self-control, but it would be too much to expect. And much of psychiatric training aims at developing just these qualities—out of an underlying stability and flexibility of personality organization, adumbrated by such terms as maturity and emotional stability. After learning to deal with disturbed patients, the psychiatrist can be expected to have the "inner confidence, objectivity and humor . . . (the ability) to remain calm and friendly and not react in a hurtful or punitive way" of which Redlich speaks.

A few points are connected in an obvious way with the effects of training: having good technical training and knowledge, preferring dynamic concepts, and perhaps developing a sharpened observational ability. The connection is somewhat less close for therapeutic optimism. An argument could be made that this is a facet of a more basic character attitude. Nevertheless, Menninger mentions "a faith in the inherent capacity of the individual, even the sickest individual, for reconstruction . . . confidence in the inherent revival powers of the

human personality" as part of "a certain philosophy, a point of view, a scale of values" for the inculcation and development of which psychiatric education must provide.

Finally, there are the three points made by Ruesch, none of which has been mentioned by the other psychiatrists. Ruesch's intention was slightly different, however. He was trying to find characteristics in the life histories of psychiatrists that would differentiate types of specialty rather than degrees of competence. (It should be added that the basis of his conclusions in data was sketchy, at best.) He concluded that general psychiatrists and those who were not primarily psychotherapists were more likely to have experienced severe traumas in childhood and to have come from socially upward mobile families, while therapists were more likely to have experienced "culture change" (mainly moving from Europe to America) or "bi-cultural experience in early childhood."

A few special words should be said on the trainability of the various qualities mentioned. It was raised specifically as an issue only in Menninger's article, "What Are the Goals of Psychiatric Education?" Here he says, in brief, that the goal is to produce a certain kind of man, a good psychiatrist, and he describes in some detail the qualities that the well-trained psychiatrist should have:

. . . integrity, intelligence, sensitivity to human suffering and numerous other such attributes . . . which antedate all medical training. The psychiatrist must be, in the platonic sense, a *good* man; he is by definition a scientist, committed to open-minded but critical search for empirical truth. Similarly, he is, *a priori*, a physician, dedicated to the care and treatment of sick people. . . . he should be a man of broad cultural interests and background, conversant with the principal social, scientific, literary and artistic developments of his own time and of the past. But if a man is not all of these things by the time he has entered medical school, it is probably too late for him to acquire them. There are exceptions, of course, and in all instances the psychiatric training will continue to put emphasis upon them and give opportunity for their development [1952, p. 153].

And again:

The psychiatrist as a person is more important than the psychiatrist as a technician or scientist. What he *is* has more effect upon his patients than anything he *does*. Because of the intimate relationship between patient and psychiatrist, the value systems, standards, in-

terests and ideals of the doctor become very important. As we have said above, these are for the most part characteristics of the man before he has even gone to medical school, and consequently, selection of those men presumably capable of becoming good psychiatrists is a responsibility of the school. Even those who have been selected, however, should be reminded frequently of this elementary fact, that their effectiveness as therapists depends in large part upon the stature and breadth of their own personalities [p. 156].

Though some colleagues might differ with Menninger on some of the specific points described as necessarily present before psychiatric training, the general point of view seems to be widely accepted. Indeed, the whole enterprise of screening applicants for psychiatric residencies on the basis of a personality assessment presupposes that it is either impossible or uneconomic to try to develop certain personal qualities in a training program.

A psychiatric resident is, nevertheless, expected to develop as a person. If he has within him only the germs of certain necessary qualities, they must exist in a flexible personality that is oriented toward further growth. This is one of the points on which there was most agreement; it was made by three men in describing the qualities needed in applicants and by one in describing the effective psychiatrist. Murray said that the good psychiatrist should have "Flexibility: ability to absorb other points of view, to change, to develop mentally." Bowman had some interesting comments on this issue:

For thirty years now I have been engaged in the selection of residents and have been observing the progress of these men. I must confess at the end of thirty years I am much more skeptical of anyone's ability to pick out the best men and predict what will happen to them after they get out. Time and time again I have seen very bright young men from our very best medical schools who look like world-beaters, who somehow lose something of the vital spark shortly after they finish residency training. I can think of several men who, during their first year of residency at the Boston Psychopathic, seemed among the very poorest men we had who are now professors of psychiatry, who are doing extremely well, and whose success I would never have predicted. There is a quality which I think we have not yet learned to measure in any way. This is the individual's capacity for growth and development. I see so many men who seem mediocre in the start; they go ahead, however, and develop steadily and keep on growing with the years, finally becoming outstanding men in the profession. I see others who start out as world-beaters and who sort of give up shortly.

Others are less skeptical of their ability to assess this elusive quality. Smith uses the device of looking for a "history of progressive growth of interest in training that shows a definite goal drive and consistency," and this technique of finding evidence through the interview that the person has been recently in the process of growth is probably used more than any other.

#### DIFFERENCES BETWEEN PSYCHIATRISTS, PSYCHOTHERAPISTS AND PSYCHOANALYSTS

A tabulation of expert opinion on personality requisites for the practice of psychotherapy and psychoanalysis was made, similar to the one for psychiatrists in Table I-15.1; it is found, with some discussion, in Appendix 15.1 (Table 15.2). Here we are concerned with it mainly to find out the extent to which expert opinion delineates different types of personality as suitable for these three overlapping professions.

To begin with, a number of apparent differences must be attributed to artifacts of the data. There are many more sources of opinion on psychoanalysts than on the other two professions put together; therefore, it should not be surprising to find that quite a number of qualities were mentioned in discussions of psychoanalysts only. Second, the questions involved were not by any means the same. Some discussions were frankly Utopian in orientation, seeking to depict a professional ideal; others were reports of specific investigations of personality qualities found in practitioners or in apparently acceptable applicants. In reference to negative qualities, there are a number of general statements and then (for psychoanalysis) a large number of replies to two parts of Dr. R. P. Knight's questionnaire investigations. In one, he asked training analysts for types of psychopathology that would be considered disqualifying, and in the other he collected characterizations of psychoanalytic candidates who had been dropped from training. (See Appendix 15.3.)

It would be a great deal easier to make sense of the agreements and differences if all of the data had been gathered in a uniform fashion, preferably being oriented toward a description of necessary (rather than merely desirable) qualities and disqualifying (rather than merely undesirable) ones. When inquiries are made with open-ended questions, or when one collects independent essays on the topic, there are bound to be many differences due simply to idiosyn-

crasies of vocabulary, the occurrence of pet concepts, and the fact that it is extremely difficult to be comprehensive unless some kind of guide or checklist is used to remind one of considerations that may be omitted inadvertently.

When all allowances of this kind have been made, there still seem to be some differences of emphasis in Tables I-15.1 and 15.2. Psychiatry does overlap with both psychotherapy and psychoanalysis and, at least in some definitions, includes them both, but it certainly also includes work functions that are foreign to the other two. *Ability to work with an institutional team, leadership ability*, and perhaps *speed of work* are very directly related to the special demands that are made on a hospital psychiatrist outside his role as a therapist. Of necessity, he must make professional use of more superficial, less prolonged, and less intense relationships with people, wherefore it seems reasonable that *sense of humor* and *likeability* should be mentioned only in discussions of psychiatry. Being involved in administration and therapeutic management, his role is inevitably more active than that of the psychotherapist or psychoanalyst; thus it is not surprising that we find no mention of *capacity to tolerate the stress of the therapeutic relationship without having to act*, or *ability to feel one's own affects without having to express them*; also related may be the fact that experts do not specify *excessive therapeutic zeal* as a danger in this field, whereas they do for the others. There is less emphasis on aspects of personality having to do with deep or subtle understanding; thus, *introspectiveness, insight into one's self, creativity, sublimated voyeurism, grasp of cultural implications and of the relativity of behavior* are all mentioned in discussions of psychoanalysts and psychotherapists but not of psychiatrists.

Because the psychiatrist must at times be active and directive, people who have authoritarian personality structures and a need to dominate are perhaps not so unsuitable as they are considered to be for psychoanalysis and psychotherapy; at any rate these qualities were not mentioned as contraindications, nor were the various types of *overconcern with attitudes of others*.

Many of these differentiating qualities (together with the fact that most of the emphasis on *warmth* comes from articles on psychiatrists) connote an active, manipulative, responsibility-taking type of relationship to patients; they suggest a conception of therapy by emotional support, direct suggestion or influence, and by giving the patient a strong and effective person with whom to identify himself.

There is a miscellaneous residue of differentiating terms that do not seem to follow any consistent pattern. *Having a happy marriage and coming from an upward-mobile family* were mentioned only for psychiatrists, and the following were omitted only for them—*firmness and authority, ability to function adequately despite personal problems, interest in own growth*, and (as disqualifications) *severe neuroses, projection, too much or too little anxiety*. All these omissions are implied in other qualities of work that were mentioned, however.

Shifting the focus to the psychotherapist, we should first utter a special caveat: there are fewer discussions about the personality of the psychotherapist than about either of the other two, and they tend to be briefer, so there are perforce many more categories that fail to be mentioned for psychotherapists. We are unable to make out any consistent pattern in these omissions. Similarly, there were only three qualities mentioned for psychotherapists only, and they are all closely related to somewhat differently worded concepts in discussions of psychiatry and psychoanalysis. They are: *ability to identify oneself with the patient's need for change* (cf. *ability to identify oneself with the therapeutic process*); *overcompliant inability to say no* (cf. *over-passivity*); and "*ambivert or dilated type*" (cf. *richness, breadth of personality*).

The following qualities, which were omitted only for psychotherapists, will be listed without any comment other than that they were in some cases undoubtedly taken for granted, in others merely overlooked, and in some represent desirable but nonessential qualities that may have been omitted only because the few discussions about psychotherapists tended to be somewhat less Utopian than some of the others: *intuition, ability to win confidence, ability to remain understanding despite attacks, dynamic convictions, interest in people, concern with human problems in their universal aspects, tolerance, some success in life, physical health, age, performance in previous training, responsibility, and, as contraindications, rigid authoritarian political beliefs, psychopathic personality, addiction, and overt perversion*.

When we examine the qualities that are attributed only to psychoanalysts and those that are omitted only in regard to them, there seems to be one main basis for differentiation: differences in technique. Compared to most kinds of psychotherapy and psychiatric treatment, psychoanalysis is longer, deeper-going, and requires more

patience and more passivity from the doctor. It is probably for that reason that we find the following group of qualities mentioned only in discussions of the psychoanalyst's personality: *tact*, *fundamental pleasure in listening*, and (also possibly related) *spontaneity vs. stereotypy of manner*. Since there were so very many discussions concerning psychoanalysts, the few qualities omitted there but mentioned as requisites for psychiatry or psychotherapy take on particular significance. *General appearance and manner* may reasonably not be thought important because of the "psychoanalytic incognito" and the length of contact, making more superficial impressions unimportant. Similarly, *extroversion* and *therapeutic optimism* are not necessary for the analyst because he does not work in such a direct, actively inspirational way but is concerned much more with phantasies, dreams, and inner matters. Therefore it is not surprising that whenever the issue of extroversion came up in discussions of psychoanalysts, the emphasis was on the desirability of *introversion*, and *extreme extroversion* was even listed as one of the reasons a candidate had to be dropped from psychoanalytic training. Among the contrary indications, *overnormality* was mentioned only for psychoanalysis, and discussions of this quality make it sound similar to the type of jolly extroversion that some people think desirable in a psychiatrist or even certain kinds of psychotherapists. A *manic*, *over-reactive* personality structure, *acting out*, and *counterphobic mechanisms* are unwanted in a psychoanalyst because of the nature of the technique. On the other hand, these aspects of technique may tempt *parasitic* or *rigid* people to fall into an opposite kind of technical error. (For further discussion see Appendix 15.4.)

The many items in our table that are found only in the discussion on psychoanalysts can be accounted for by differences in vocabulary and by the fact that there were many more sources available. It should not be surprising that training analysts, writing for each other, would use certain psychoanalytic concepts not referred to as such by writers discussing psychiatric selection. Thus, both psychoanalysts and psychiatrists are supposed to have *curiosity* and *compassion*, but only in the literature on psychoanalysts is it specified that the candidate's motives should be *sublimated* and that *seeking direct instinctual gratification in work* is a disqualification.

To sum up: the degree of overlap in discussions of the personalities of psychiatrists, psychoanalysts, and psychotherapists is im-

pressive. Moreover, the differences that we could ferret out are either fairly obviously related to the role of a resident or hospital psychiatrist in administration and in therapeutic management of the patient or else are specific to types of treatment that were not held in much esteem at the Menninger School of Psychiatry. The psychiatrist of today can hardly function effectively if he is not a skilled psychotherapist, and the psychotherapy taught at the Menninger School of Psychiatry, and in many other psychiatric training centers as well, is for the most part psychoanalytically oriented.

Statistical listings such as the tables to this chapter appear bare indeed when one compares them with the rich and vividly detailed description in many of our sources. For economy of space, further discussions of the variables tabulated, with quotations from our sources, have been omitted here but may be found in Appendix 15.1.

#### THE ISSUE OF SPECIFICITY TO PSYCHIATRY

In some discussions of "the good psychiatrist," one reads descriptions of the kind of man whom one might hope to encounter in any profession and who might be expected to do well in almost any type of work. We have already shown that there is relatively little specificity to the particular disciplines of psychiatry, psychotherapy, and psychoanalysis in the opinions collated here; may it not be that these qualities would apply just as well to various other occupational groups?

There are a couple of indications, on the level of opinion, that little differentiation can be made between the personality of "the good psychiatrist" and that of "the good clinical psychologist." A committee of the American Psychological Association drew up a list of "qualities of the good clinical psychologist" (Shakow *et al.*, 1947); every one of them can be fitted immediately into Table I-15.1 or Table 15.2:

1. Superior intellectual ability.
2. Originality, resourcefulness and versatility.
3. "Fresh and insatiable" curiosity; "self-learner."
4. Interest in persons as individuals rather than material for manipulation—a regard for the integrity of other persons.
5. Insight into own personality characteristics; sense of humor.
6. Sensitivity to the complexities of motivation.

7. Tolerance; "unarrogance."
8. Ability to adopt a "therapeutic" attitude; ability to establish warm and effective relationships with others.
9. Industry; methodical work habits; ability to tolerate pressure.
10. Acceptance of responsibility.
11. Tact and cooperativeness.
12. Integrity, self-control and stability.
13. Discriminating sense of ethical values.
14. Breadth of cultural background—"educated man."
15. Deep interest in psychology, especially in its clinical aspects.

In the research on the selection of clinical psychologists, Kelly and Fiske had their judges rate the subjects on a list of 32 variables of personality. When these ratings were correlated with subsequent criterion measures of actual performance, six variables showed a significant degree of relationship with performance. These, too, may easily be located in our tables: broad interests; adaptability and flexibility; imagination; motivation for scientific understanding of people; insight into himself; and quality of intellectual accomplishments.

Aside from the fact that there are many variables in our extensive lists that are not directly touched on, the qualities conducive to success in clinical psychology seem to be remarkably similar to those needed for psychiatry, psychotherapy, or psychoanalysis. This should hardly be surprising since all four types of work share the functions of diagnosis and psychotherapy.

Similarly, comparisons could be made to show that the qualities of personality thought to be desirable for social work or general medicine would overlap extensively with the ones we have been discussing in this chapter. The more dissimilar the work, the more one would expect the lists of requisites to differ.

Perhaps it is impossible to describe a professional ideal without describing to some extent a good man, a generally superior human being. As John Stuart Mill wrote, "Men are men before they are lawyers or physicians or manufacturers; and if you make them capable and sensible men, they will make themselves capable and sensible lawyers or physicians." It may well be that a good psychiatrist does have much of what it would take to be a good lawyer, a manufacturer, or even a creative artist. The mention of the last field reminds us that there is such a thing as specific talent for certain types of work but that such talent alone is not enough. A man

may have enormous gifts in the graphic arts, but if he lacks the maturity and the self-discipline to do any sustained work, he will never succeed in his field.

The reminder that there are specific talents for certain types of work also raises the question: Is psychiatry one of these types of work? Is there a specifically psychiatric talent? There is nothing in the present survey of expert opinion that enables us to distinguish between general qualities that pertain to success and failure in almost any vocation and those that may be specific to psychiatry. Yet there is still a difference. Although our list consists in the main of qualities that would doubtless be useful to human beings in almost any endeavor, according to our experts many of them are *indispensable* to honest and effective work in psychiatry.

From a practical standpoint, it makes little difference that we do not know the qualities that are uniquely essential to psychiatry. If we can get some greater degree of certainty about which aspects of personality to look for when selecting potential psychiatrists, we shall be quite content, whether or not these same qualities might also be found in good prospective ministers or merchants.

TABLE I-15.1. SUMMARY OF EXPERT OPINION ON  
PERSONALITY REQUISITES FOR PSYCHIATRY \*

QUALITIES TO BE SOUGHT IN APPLICANTS FOR PSYCHIATRIC TRAINING	QUALITIES CHARACTERIZING THE TRAINED (EFFECTIVE) PSYCHIATRIST
<i>I. Abilities and Capacities</i>	
	A. Intellectual
1. Superior intelligence (MSP; Menninger; Hastings)	1. Intelligence (Redlich; Murray) 2. Common sense (Redlich) 3. Observational ability (Murray) 4. Imagination (Murray)
	B. Interpersonal: Receptive
1. Intuitiveness (Greenhill) 2. Capacity for understanding (Grinker; Waggoner) 3. Empathy (MSP; Greenhill) 4. Psychological-mindedness (MSP)	1. Intuition (Murray) 2. Sensitivity to subtle dynamics of human behavior (Murray) 3. Empathy (Murray)

\* Where a corresponding quality is lacking in a column, the number is omitted.

TABLE I-15.1 (Cont.)

QUALITIES TO BE SOUGHT IN APPLICANTS FOR PSYCHIATRIC TRAINING	QUALITIES CHARACTERIZING THE TRAINED (EFFECTIVE) PSYCHIATRIST
<b>C. Interpersonal: Interactive and Relational</b>	
1. Verbal facility (MSP)	
2. Capacity to attract friendship (Levine)	2. Likeability: ability to win affection (Murray)
3. Ability to interrelate with many types of people (Smith)	4. Ability to win respect and trust (Murray)
5. Ability to work harmoniously with institutional colleagues (Levine; Noyes)	5. Ability to work as member of a team; absence of annoying traits (Murray)
6. Leadership ability (Smith)	6. Ability to raise morale, maintain serene atmosphere (Murray)
<b>D. Inner-personal</b>	
5. Sense of humor (Bellinger; Levine)	1. Self-control (Redlich)
6. Relative freedom and ease in reacting (Waggoner)	2. Objectivity and detachment (Redlich; Murray)
7. Capacity for growth and development (Bowman; Grinker; Smith)	3. Ability to tolerate strains, frustrations (Murray)
	4. Ability to remain understanding in the face of personal attack (Menninger)
	5. Sense of humor (Redlich)
	7. Flexibility: Ability to change, develop (Murray)
<b>E. Miscellaneous</b>	
1. Initiative (Noyes)	1. General energy level, zest, initiative (Murray)
	2. Speed and perseverance in work (Murray)

TABLE I-15.1 (Cont.)

QUALITIES TO BE SOUGHT IN  
APPLICANTS FOR PSYCHIATRIC  
TRAINING

QUALITIES CHARACTERIZING THE  
TRAINED (EFFECTIVE)  
PSYCHIATRIST

II. *Attitudes, Interests, and Values*

A. General Characteristics of Attitudes and Values

1. Breadth of (nonmedical) interests (Ewalt; Levine)
1. Breadth of interests (Murray)

B. Interest in Psychiatry

1. Interest in subject-matter (Murray)
2. Preference for dynamic concepts (Murray)

C. Attitudes toward Patients and People Generally

1. Interest in people (Bellinger; Waggoner)
2. Concern with human problems in their universal aspects (MSP)
3. Respect for the dignity and integrity of the individual (Menninger)
4. Tolerance (MSP)
5. Sense of social responsibility (Menninger)
6. Therapeutic optimism (Murray; Menninger)
7. (*Negative quality*) Communist beliefs (Bellinger)

III. *Motives*

A. Specific Motives for Specializing in Psychiatry

1. Desire to help relieve suffering (MSP; Menninger)
2. Need to search for empirical truth (Menninger)

TABLE I-15.1 (Cont.)

QUALITIES TO BE SOUGHT IN  
APPLICANTS FOR PSYCHIATRIC  
TRAINING

QUALITIES CHARACTERIZING THE  
TRAINED (EFFECTIVE)  
PSYCHIATRIST

*IV. Life Experiences and Achievements*

A. Childhood and Adolescence

1. Family background of upward social mobility (Ruesch)
2. Childhood traumas (among those in administration, teaching, and research) (Ruesch)
3. Experiences of culture change and contact with other cultures (Ruesch)

B. General Features of Adult Experience

- |   |  |
|---|--|
| 1. Breadth and variety of life experience (Ewalt) | 2. Happy marriage, with children if possible (Redlich) |
| 2. Happy marriage (Ewalt)                         |  |
| 3. Satisfactory accomplishments (Waggoner)        |  |
| 4. Physical health (MSP; Bellinger)               |  |
| 5. Age (Bellinger)                                |  |

C. Intellectual Achievements

- |  |  |
|--|--|
| 1. Cultural background (MSP)   | 1. Broad cultural interests, viewpoint (Menninger; Redlich)      |
| 2. Good scholastic record in medical school (Bellinger; Ewalt; Bowman; Hastings). <i>Qualifications</i> by Bellinger; Ewalt; Gildea) | 2. Ability to learn psychiatry; marks in medical school (Murray) |
| 2a. Graduation from Class A medical school (Bellinger)   |  |
| 3. Performance in internship, psychiatric clerkship (Bowman; Gildea; Levine)   |  |
|  | 4. Good technical training (Murray; Menninger; Redlich)          |

TABLE I-15.1 (Cont.)

QUALITIES TO BE SOUGHT IN APPLICANTS FOR PSYCHIATRIC TRAINING	QUALITIES CHARACTERIZING THE TRAINED (EFFECTIVE) PSYCHIATRIST
---	---

V. *Other Traits of Personality*

A. Adjustment and Health

- |   |                                 |
|---|---------------------------------|
| 1. Maturity (Levine; Smith)                     | 2. Emotional stability (Murray) |
| 2. Emotional stability (MSP)                    |                                 |
| 3. Relative freedom from symptoms<br>(Waggoner) |                                 |

A'. *Psychopathology (Negative Indications)*

- |   |   |
|---|---|
| 1. <i>Addiction</i> (Hastings)                      | 1. <i>Alcoholism</i> (Murray)                                 |
| 2. <i>Overt homosexuality</i> (Hastings;<br>Levine) |   |
| 3. <i>Psychopathic tendencies</i><br>(Bellinger)    | 3. <i>Sleeping with nurses in the hos-<br/>pital</i> (Murray) |
|   | 4. <i>Egocentric traits</i> (Murray)                          |

B. General Evaluative Traits

- |  |  |
|--|--|
| 1. Integrity of character (MSP;<br>Smith; Menninger)   | 1. Integrity; truthfulness (Murray)          |
| 2. Sincerity (Waggoner)                                |  |
| 3. Stature and breadth of person-<br>ality (Menninger) |  |
| 4. Acceptance of responsibility<br>(Levine)            | 4. Dependability (Murray)                    |
| 5. General appearance and manner<br>(Bellinger)        | 5. General appearance and manner<br>(Murray) |

C. Emotional and Interpersonal Traits

- |   |  |
|---|--|
| 1. Emotional warmth (MSP;<br>Levine; Bellinger) | 1. Warmth (Redlich; Murray)                          |
|   | 2. Sympathy, kindness, consider-<br>ateness (Murray) |
|   | 3. Co-operativeness (Murray)                         |

D. Inner-personal Traits

- |  |                               |
|--|-------------------------------|
| 1. Independence without hostility<br>to authority (Levine) | 2. Inner confidence (Redlich) |
| 3. Extraversion (Bellinger)                                |                               |

## PERSONALITY DIFFERENCES BETWEEN BETTER AND POORER RESIDENTS

TO MAKE AN EMPIRICAL STUDY OF THE PERSONALITY TRAITS THAT ARE associated with good performance in psychiatric residency, it is necessary to have a reliable measure of such performance (which we have in the Supervisors' Evaluations of competence) and also valid measures of the qualities of personality. Measures of the latter kind are not easy to come by; there are no tests, no standardized interview procedures, or the like that lie ready for use, their validity proved by satisfactory evidence. Instead, one has to rely once again on the judgment of trained observers. Moreover, the judges of personality traits must have extensive knowledge of the persons to be rated and must be trained in making the kinds of inference needed. Two ways of measuring personality are generally accepted as having the greatest intrinsic validity: (1) pooled judgments by raters who have studied the subjects through interviews, life history materials, psychological tests, personal contact, etc.; and (2) judgments by clinicians who have had opportunities to observe the types of behavior in question directly and over a long period of time (as may be done in long-term psychotherapy or psychoanalysis).

Fortunately, we have two sources of information about our final group of subjects which correspond rather closely to each of these two kinds of personality criteria. Let us see what both of these types

of data tell about the personality of the psychiatric resident and how well they agree.

#### PERSONALITY VARIABLES AS PREDICTORS

In the Predictive Study two judges went through the entire file of assessment data on each of sixty-four applicants, analyzing them in great detail and making ratings on thirty-two personality variables. When these two sets of independent ratings are averaged, they make a good approximation to the first type of criterion described above. They have the advantage of being quantitative measures, applied to all subjects of Classes V and VI *at the time of their application* and describing their personalities as they were at that time.\* The last point is crucial if our interest is not just theoretical or academic; the assessor should know which qualities will be evident initially and which he can expect to find in the later stages of a man's training. Granted, the strains of being scrutinized will cause a man to become defensive and add to the difficulties of getting valid measures. If personality variables can predict competence in psychiatric work anyway, they should be the more valuable.

Details about the ratings and the procedure followed are given in Appendix 16.1; there, too, may be found Table 16.1 with the full correlational results. There were three variables which correlated to a surprisingly high degree (.40 or better) with Supervisors' Evaluations of Over-all Competence: Genuineness *vs.* Façade, Social Adjustment with Co-workers, and Freedom from Status-Mindedness. That is, the combined judgments of Genuineness made by two psychologists, who had reviewed all tests, interviews, credentials, etc. on the sixty-four subjects and judged these qualities, predicted quite well which applicants would be highly rated as residents; and their pooled judgments on each of the other two variables (Social Adjustment, Status-Mindedness) did equally well. These three variables have in common a bearing on effectiveness in human relationships. Also, unlike many of our other variables (about which the same thing could be said), they lie too close to the surface to be concealed;

\* Note that the predictive ratings were aimed forward in time—the judges tried to estimate the level of performance expected from a subject in his (future) residency work—but the personality ratings summed up the picture of the subject as he appeared during assessment at the time of application (see Chap. 10).

and, as compared to possibly subtler aspects of personality such as Psychological-Mindedness and Empathy, they are less likely to exist in chrysalis (not fully called upon and thus not made manifest before the applicant has undertaken psychiatric work). Thus, conceivably, other variables may be more intrinsically related to psychiatric work yet, being harder to assess, yield lower correlations.

A second group of variables all correlate above .30 with Over-all Competence. Three of them may be looked on as indexes of general adequacy of functioning, or psychological health: Self-Objectivity, Mature Heterosexual Adjustment, and Adequate Emotional Control. What we called Self-Objectivity is very close to what was referred to by most of the sources in the preceding chapter as self-insight, and its relevance to psychiatric work needs no comment. There is surely a less direct relationship between sexual adjustment and psychiatric work, however. Accordingly, we rated not so much the level of maturity in sexual relationships that the subject had actually attained as his orientation, the direction in which he seemed to be moving. It was a measure of a man's capacity for intimacy and his capacity to love—which may very well be the best index of his readiness for the close and loving (yet objective) relationship of psychiatric care. Our measure of emotional control was deliberately defined with its high point an optimum; by Adequate Emotional Control we meant the capacity to release as well as to restrain and the general effectiveness of defenses. Even more than the other two, this characteristic is a barometer of general emotional health.

Another cluster within this second group of highly correlated variables may have in common a kindly and understanding approach to others. They are: the Need to Help, Empathy, and Consciousness of Social Injustice. The capacity to be warm and loving, implied in Adequate Emotional Control and Mature Heterosexual Adjustment, is directly stated in the Need to Help, which emphasizes the motivational aspect. Consciousness of Social Injustice may be looked on as the attitudinal or ideological expression of this trend.

Next (still significantly related to competence at better than the .01 point), there are two variables we have described earlier as being special aspects of Adequate Emotional Control: Emotional Appropriateness and Adequate Handling of Hostility; and a third, Internalized Interest in Psychiatric Work, which speaks for itself. Like Empathy (to which Emotional Appropriateness is closely related), the former pair of variables are subtle qualities, not easy to assess

since they have to do with delicate emotional balances. This fact may account for their being somewhat less correlated with the criterion than the other variables so far mentioned.

There are ten more variables correlated with Over-all Competence with coefficients ranging from .21 to .28. These correlations are significant but small, yet it is tempting to argue that in several instances the lower correlations are due to errors in the measurement of elusive qualities, not to their being actually less important in psychiatric work: Psychological-Mindedness and Capacity for Personality Growth especially. These are two qualities about which we have the impression that they were least fully manifest at the time of application. Objectivity toward Authority Figures and Freedom from Ethnocentric Prejudice, two closely related traits in the authoritarian personality syndrome, according to Adorno *et al.* (1950), indicate the relative incompatibility of that personality type with the kind of psychiatry taught at the Menninger School of Psychiatry. In addition to its more general implications for personality organization, Objectivity toward Authority Figures is directly involved in satisfactory work as a psychiatric resident under supervision in an institutional structure. It is especially interesting to find a significant relationship for Freedom from Ethnocentric Prejudice, since as we have seen (Chaps. 4 and 13), it was extremely rare to find it markedly present in an applicant, and questionnaire measures had no validity as predictors of competence. Objectivity toward Dependent Figures is also essentially a type of orientation toward others; all three of these last-mentioned variables are measures of objectivity.

Two measures of attitudes toward the self show slight but reliable predictive powers: Self-Confidence and Feelings of Security. It may be recalled that, in Chap. 14, one of the judges found he had erred in attributing to these attitudes more relevance to the criterion than they have. And two measures of ability also show some relationship: Clarity of Thought and Quality of Verbalization. Cultural Wealth was related to Over-all Competence perhaps because it implies a generally sensitive person. Freedom from Material Pre-occupations correlated with competence because it is a measure of undesirable extrinsic motivations.

The remaining handful of variables may be divided into two groups: the few that may be slightly related to the criterion in the scoring of only one judge or the other, and the few that are probably irrelevant. Freedom from Stereotypy of Thought is tangentially re-

lated to the intellectual aspects of psychiatry in the same way as clarity but to a slighter degree. And Well vs. Poorly Internalized Ethical Standards could not be rated at all for some applicants since there was so little evidence bearing on it. Important though it undoubtedly is to exclude persons who are definitely unethical, this variable is hard to scale, and what few distinctions could be made did not seem very important. The other group, which showed no tendency toward significant validity, includes Sense of Humor, Perceptual Sensitivity, and Psychological Curiosity. (See below for discussion of these variables. A few other variables could not be correlated, not having been rated on enough cases.)

#### PERSONALITY DIFFERENCES BETWEEN CRITERION GROUPS

Let us look next at a set of judgments made by skilled clinicians about the personalities of psychiatric residents. Although not quantitative, these judgments have the advantage of being based on extensive observation of the very kind of behavior in which we are most interested. They thus comprise a usable approximation to the second type of personality criterion mentioned above.

The psychiatrists who supervise a resident's work are in the best position to make observations of the relevant aspects of his personality in operation. Their comments have been classified and tallied for the 33 residents who were given the highest ratings on Overall Competence (Highs) and the 33 who were given the lowest ratings (Lows)—the top and bottom 13 per cent of the 247 psychiatrists trained in the Menninger School of Psychiatry during 1946–51.\* (Full details on the method used and complete tabulations are given in Appendix 16.2 and Table 16.2.)

A slight inequality affects the comparison of the two groups: more statements were made about the Highs than about the Lows, since the Highs stayed an average of twenty-five months in residency and the Lows only twenty months. Thus, more supervisors (on the average, two more) had a chance to become familiar with the work of the Highs. But even if we take into account the length of stay in

\* Virtually the same groups of residents were used for ascertaining the personality qualities of psychotherapists in earlier reports (Luborsky, 1952a, b). The correlation between Over-all Competence and Psychotherapeutic Competence is so high (.86 for Classes V–VI) that the extreme groups on either criterion are almost identical.

the School, we find proportionately more statements about the Highs. Apparently the less successful were slightly less well-known to their teachers.

We also searched through considerable additional data to find further evidence relating to the qualities emphasized by the supervisors—especially the qualities that markedly distinguished the two groups. These data included the test reports written by a psychologist \* for the Admissions Committee at the time of the residents' application for training, some scores from the tests, background information, opinions expressed by patients about these residents, the psychotherapy process notes written by the residents, and information describing the types of practice they set up after leaving Topeka.

In the following section we shall take up these qualities in six groupings, discussing especially the ones that were most discriminating and looking into the collateral data for confirmation.

#### WORK CAPACITIES

There are few surprises in this first group; one would expect supervisors to describe men they had rated at opposite extremes as having very different work-relevant abilities. The better residents were seen significantly more often as intelligent, responsible, having good judgment, clear in thought and verbal expression, sensitive and empathic, much interested in psychiatry, creative and imaginative, self-reliant, honest and knowledgeable about psychiatry. The groups were least differentiated by the term conscientiousness. It was, in fact, the positive quality most frequently attributed to Highs and Lows alike.

We happen to have independent supporting evidence for some of these differences—which, of course, says nothing about the truth or falsity of those that are unsupported. On the Wechsler-Bellevue test, the Total IQ's of the Highs were slightly higher, and the Verbal IQ's were significantly higher than those of the Lows. The psycho-

\* Only the test reports written by *one* psychologist (on 23 Highs and 21 Lows) were included to insure that the reports would as much as possible be comparable in content. The comparison of the test report qualities with the ones mentioned by supervisors is not entirely fair to the tests. The test reports were very brief—about two hundred words—they did not follow any set outline, and only a few of the qualities were mentioned with any regularity. Either good or poor endowment was mentioned in 31 cases. Next in order of frequency are the following: obsessiveness and circumstantiality, ability to make rapport, passivity, empathy, and intellectualization.

logical test reports gave further evidence of a difference in intelligence in a broader sense: *endowment* was the most differentiating of the qualities mentioned.\* The fact that knowledge of psychiatry and related fields did not markedly distinguish the two groups is confirmed by the very low positive correlation between a test of academic knowledge given by the School of Psychiatry and the supervisors' ratings of the residents' competence (Chap. 8).

The relevance of self-reliance and self-confidence appears to be supported by the results of a self-concept questionnaire (Mayman, 1953), which was given to small groups of these Highs and Lows plus some residents of intermediate capacities. The three groups fell in the expected order (High, Middle, Low) in the certainty of their self-descriptions and good opinions of themselves.

Not only was the interest in psychiatry of Highs greater than of Lows but it was significantly more focalized. In the Supplementary Face Sheet questionnaire, applicants were asked to check the kinds of psychiatric work functions in which they would like to spend time. The Lows checked significantly more than the Highs.

The Highs have published significantly more scientific papers than the Lows. Although we did not undertake the delicate task of judging the quality of publications, this quantitative finding may be evidence supporting the observation of the supervisors on creativity and imaginativeness.

#### RELATIONSHIPS WITH PEOPLE

Here, too, we find that the most differentiating descriptions refer directly to behavior that would cause a man to be rated highly: good relationships with patients, good relationships with co-workers, and good relationships with supervisors. By contrast, ward personnel often reported disruption of their work by Lows. Consistent with these good relationships for the Highs was their being seen as less defensive and more likeable to supervisors (the latter being a trend of borderline significance). The Highs' greater willingness and ability to learn and less need for supervision—see below—are probably relevant in this connection also. This does not mean that Highs took

\* The psychologist supplied the following description of the bases for his judgment of a resident's endowment in addition to IQ: information score, number of good M's, quality of verbalization, good concepts (e.g., in the Rorschach, a well-crystallized response such as "Voltaire"), subtlety of formulation, awareness of multilayeredness of life and phenomena. "It could be put most simply in Lewinian terms—breadth and differentiation of life space," he wrote.

all advice uncritically or that they did not often disagree with their supervisors. By contrast, the Lows more often tended to become passively aggressive, grumbled or broke rules or were obtuse in understanding the supervisors' views.

Psychiatrists are generally expected to excel in the area of relationships with people. It is amply clear that the Highs were better able to get people to like and respect them (or, that people who were liked and respected were thought to be competent psychiatric residents). The point can be seen in another kind of data: the generally high correlation of Liking with other criterion variables for both Supervisors' and Peers' Evaluations (see Chaps. 8 and 13). In the Michigan Project, Supervisors' Liking also correlated highly with other criteria (Kelly and Fiske, 1951, p. 85). Supportive evidence comes from the test reports, where rapport-making ability was second only to endowment in being the most differentiating of the characteristics mentioned.

#### REGULATION OF AFFECT AND IMPULSE

A large number of qualities were grouped under this heading. Only one of them was as discriminating as those in the areas just discussed (work competence and relationships with people): self-containedness or even-temperedness. The most frequently mentioned was warmth, but the modal quality of warmth in the better group came closer to matter-of-factness than to popular misconceptions such as hand-holding, gladhanding, or hearty warmth. Appropriate warmth was described, rather than a steady glow. Two of the other significant qualities, self-containedness and objectivity, are consistent with restrained, rather than outpouring, warmth. Although the Highs were more restrained in their expressions of affect, they were nevertheless described as more energetic. The fact that intellectualization was mentioned about equally often for both groups suggests that the Highs' restraint was not due to their relying more upon isolation.

More is at issue than differences in modulation and degree of expressiveness, however. The Lows' expressions were more inappropriate to the situation. In light of the deficiencies of the Lows in sensitivity and judgment, inappropriateness is to be expected. Furthermore, when anxiety was aroused in the Highs, it was more easily bound by control mechanisms that were less immediately disruptive to the work of treatment.

The Lows found it more difficult to be sensitive to what a patient might be doing and more often dealt with it by authoritarian directiveness (as compared to the tendency of the Highs to be more appropriately firm). For example, the psychotherapy process notes showed that the Lows more readily jumped in with unnecessary reassurance, praise or blame, directive questions, opening remarks in each session, the suggestion of topics to talk about, and one-sided decisions on termination. They more frequently became angry at the patient or tried too hard to keep him comfortable.

There were overcontrolled and undercontrolled varieties of this characteristically inappropriate expressiveness of the Lows. The overcontrolled group seemed relatively colorless, flat, and brittle; such men attempted to be self-controlled but might explode into expressiveness at the wrong times. The undercontrolled group's persistent, exuberant expressiveness quite easily came to the attention of supervisors, patients, and everyone else as inappropriate.

The test reports were consistent with the supervisors' descriptions in attributing evidences of warmth slightly more often to Highs than to Lows. Also, the general conclusion about the therapist's tolerance of the patient's productions is well in line with evidence from the small group of cases for which we were able to score Fiedler's so-called "abreaction index" (see Kelly and Fiske, 1951). This index was scored from the residents' psychotherapy process notes written after each session with a patient. The scoring gives special weight to expressions of feeling by the patient about himself, the therapist, people who are close to him, and sharing of intimate details. According to this index, the patients of Highs discussed such emotion-laden and meaningful material more often than the patients of Lows.

#### MENTAL HEALTH

The supervisors described the Highs as being much better off in insight into themselves, stability, and maturity. Using one term or another, they identified about half of the Lows as severely "sick" people, and only one of the Highs as such (though several Highs were called unstable). We find partial confirmation of these descriptions in the greater number of Lows who suffered from personality problems for which they sought treatment. One required hospital in-patient care, and two had to leave residency in order to get treatment, as compared to none of the Highs.

The test reports showed these trends but to much lesser extent

than Supervisors' Evaluations. Slightly more Lows were described as prone to depression and low mood, but there were no differences in the numbers who were described as schizoid (*cf.* the supervisors' nondifferentiating term, "withdrawn") or even as having projective or paranoid tendencies. Even though these are extreme groups, none of the conventional scores of the Rorschach discriminated.

The life history data are also consistent with the points made above. More of the Highs were married, and somewhat more of the Lows were divorced, had never married, or had severe marital problems (*cf.* Chap. 13). Although, as mentioned, the Highs were more often called mature, they were significantly younger than the Lows. This does not deny that psychiatrists in general increase in skill with age and experience; it merely points out that the age when a man enters training is related to his skill. The existence of a few peculiar people in psychiatry, who, though small in numbers, have been high in public visibility has given rise to the misconception that the profession is filled with them. Our findings support the conclusion that the great bulk of competent residents (and probably of good practitioners generally) are healthy, conventionally adjusted people, whereas those who are relatively "abnormal" are concentrated on the lower fringe of competence.

#### ABILITY TO DEVELOP

Of course the qualities grouped under this heading were disproportionately attributed to the Highs. They did not as often repeat the same mistakes; they were flexible and not so defensive that they could not learn. Willingness and ability to learn differentiated Highs and Lows very clearly. The supervisors had more faith in the ultimate development of the Highs and emphasized in almost every case their good potentialities. The Lows were said to improve occasionally, but not enough to lift them out of the Low category.

It is consistent with these descriptions, then, that when the Highs needed personal treatment they applied for it (and got it) first. This was true not only of training analyses but also of therapeutic analyses and psychotherapy (see Chap. 5). Of the Highs, twenty-two had started psychoanalytic training in comparison with six of the Lows (as of the time of our follow-up questionnaires, approximately three years after residency). Although somewhat more of the Highs than the Lows said that the positions they held at the time of the last follow-up questionnaire satisfied their professional needs, the

better ones showed no inclination to stop at this level. They had many more unfulfilled desires; e.g., to be able to do more psychotherapy, more teaching, etc. It may be they were more vocal about their likes and dislikes, but these results are consistent with the idea that they did not so easily get in a rut or stay there.

Despite the fact that supervisors saw little to comment on in the Lows by way of development, this quality did not differentiate the groups very significantly, and our impression from unsystematic follow-up in recent years is that some of these men who did poorly in residency slowly blossomed after their training years were over (see Chap. 5).

Since the supervisors were describing men whom they themselves had rated as unusually good or poor residents, one would expect them to discuss these men qualitatively in ways that are consistent. But not all favorable terms were used more often to characterize Highs; some terms of generally negative connotation (such as passive and inhibited) were applied slightly more often to them than to Lows. The result of examining the qualities that did not differentiate the groups, especially those that might have been expected to, encourages our faith in the meaningfulness of these data.

We had what must be a common misconception about the quality of warmth (*cf.* Chapter 14)—expressions of it in the Highs turned out to be more subdued than we had expected. We expected much greater differences in honesty than we found. Perhaps we should have considered that if a man was very deficient in integrity, he would not be accepted or would be quickly dropped from the School. Moreover, advances beyond an acceptable level of integrity do not easily show. Conscientiousness may be similar; almost everyone of the Highs and Lows had it, so it could not discriminate. The fact that so many of both groups were described as “insecure; anxious; ill at ease” should be probably no surprise. The training program was an arduous one, and anxieties were natural. More important than the presence of anxiety was the way it was handled. Apparently the supervisors noted how often both Highs and Lows responded to the frustrations and anxieties of training with hostility or aggression.

In thinking about such findings and the related point that members of both groups were reported equally often as having trouble in expressing or handling hostility in themselves and in patients, we should remember that these are trainees. Our super-

visors' comments come from relatively early, as well as late, months of training; they tell us something about the kinds of reactions that were universal among men learning to deal with psychiatric patients. Almost anyone who is trying to become a psychiatrist is likely to show some kindness and desire to help patients. Neither impulsive lability nor quietness and withdrawal should be considered a disqualification, for these qualities showed up almost as frequently at both extremes. (Note, however, that spontaneity, as distinguished from impulsiveness, did show a tendency to be found more often in the best residents.)

The main differences between the two groups may be summarized in a thumbnail sketch: The better residents were more intelligent, sensitive, and independent in thinking and judgment. They could be warmer but also more self-contained and even-tempered, and they expressed themselves more appropriately. Their relationships with patients and others with whom they worked were better; interest in their work and learning to do it came more easily to them. Through greater stability, maturity, and self-insight, they were self-developing people.

A listing of the common qualities in a group of people is miles away from what a good character sketch of each person could tell us. The conceptions of "good" and "poor" residents that emerge seem too neat and pretty, with nearly all good on one side and all bad on the other. Much of this has to do with the method of abstracting and averaging the qualities of the two groups. Characteristics that might have rounded out the characters of the men tended to get lost if they did not happen to fit in with group trends. We can correct this partially by summarizing the ways in which several Highs and Lows were described, with a few examples from their treatment of patients.

*Resident No. 1 (Low):*

Dr. Smith, aged thirty, light hair, average height, had been married for seven years and had two children. He made a considerable sacrifice to put himself through psychiatric training. Although he was very serious about his work, his personality limitations were always in the way. Five of his supervisors commented on his having "improved" since they saw him last; five emphasized his "conscientiousness." Four said he was "not self-reliant," three commented that he was "able to accept and use supervision"; and two disagreed. Three of the supervisors said he had "poor judgment."

It is not until we get to descriptions that are more difficult to

code that we begin to get the flavor of the man. In his first year one supervisor said, "He is a man of mediocre endowment . . . so aware of his own inadequacies in dealing with patients . . . he got too involved with a very emotionally immature psychopath and was unable to untangle himself." Another said Dr. Smith had been "rather depressed over complicated and unsatisfactory relationships with some of his patients (psychopaths and paranoid) but is increasingly aware of some of his difficulties and lately has seemed much less anxious." In the next year another supervisor wrote that "he is conscientious but his letters to relatives are poor largely because of his anxiety and inability to think clearly. He gets angry at patients who get angry at him, but reacts passively." In the same year another supervisor said that "he is too impersonal and several of his patients have told the section chief they found it difficult to talk to their doctor, feeling that he doesn't understand their problems." In the last year of his residency he was improving but not doing adequate work; he was still hard working, practical and efficient but unsure of himself and inclined to take things a little too seriously. He was described as less rigid than before, and his relationships with patients were better.

We looked over the "process notes" of his psychotherapy with three patients. Dr. Smith more often discussed the patient's lack of confidence in him than did another therapist, a High who later treated the same patient. Dr. Smith often accused the patient of being mad at him or afraid of him, and he occasionally used too much technical lingo. His steadfast, plodding, serious approach was just as described by supervisors.

*Resident No. 2 (Low):*

Dr. Walgreen was an extremely overfriendly married man, thirty-one years of age. Five out of thirteen supervisors emphasized his "conscientiousness," and five his "lack of sensitivity and empathy," although two others said he had "some sensitivity." Four called him "kindly and interested in people" though two of these added, "too much so." Four said he was "interested in psychiatry and his work"; four supervisors said he lacked objectivity, and one that he had "some objectivity." He had "good relationships with his co-workers" according to supervisors, but one described poor relations with fellow team members. On his very first service (continued treatment of psychotic patients) it was noted that working closely with patients made him uncomfortable. One supervisor said "He keeps up his ward work in a satisfactory manner. He is quite verbal but not a star in the area of psychological mindedness. In working closely with a patient he shows discomfort by speaking faster and yet not saying too much." The attending physician on the service commented, "He is extremely enthusiastic and eager for instruction but frequently accepts instruction without evidence of real thinking on his part." On another

service, toward the end of his first year of training, a supervisor commented, "A compulsive somewhat rigid resident who fits into the service well. He is resourceful and capable of utilizing management and therapeutic devices for the more acutely psychotic patients. He works closely with aides and nurses and consults readily with the section chief when in trouble. He comes out nights to see disturbed patients and meet personnel. He tends to neglect paper work and must be prodded to get it completed. He knows his patients' present status and history well. He is anxious to please and reports fully to questions asked. If he fails to know, he looks it up. It is difficult to tell how he would do in a closer doctor-patient relationship as he is too busy to form any."

In his second year he attended a psychotherapy control group. The control group leader mentioned in his evaluation, "The one thing that comes to mind immediately about him is that he has a tendency to be rather aggressive toward the presentor. It is a pressure to show that he can quickly understand the unconscious meaning of something or other—such as a slip the presentor might make which might mean that the presentor wants reassurance himself from the patient. On the asset side he is extremely interested, has lots of enthusiasm, is always eager to talk about his work and discuss the work of others."

At the end of the second year, when he had been working on a service that required close relationships with patients and team members, his difficulties increased. The supervisor said: "Dr. Walgreen has difficulty relating himself to other members of the team; he always seemed able to arouse resentment by his obsequiousness, insensitivity and indecisiveness. With patients he made many errors in judgment and could never seem to be able to interpret what they were saying to him. He was overconcerned about their progress and often stayed at the hospital overtime and all day on Sundays to be with his patients. While we felt this may reflect a desirable quality, it is to a much greater extent a reflection of his own anxiety. He was always in a habit of trying to please the patients so as not to incur their wrath, and in this way often defeated the goals of treatment. He was quite obsessive in writing abstracts and letters to relatives. When he was put much more on his own, toward the end of the service, it seemed to produce immediate results in that he became more definite and able to handle responsibility better. He certainly has done a much better job than when he first came. And, he is not the overbearing person that he was." On his next service, in the first part of his third year, his performance improved slightly in a work situation with somewhat less supervision and somewhat less close contact with patients.

Dr. Walgreen submitted copies of the process notes of the therapy of three patients that he had treated. The following happenings seem noteworthy: (1) He explicitly took the husband's side

against a wife when he could have helped the patient more by remaining neutral. (2) There was an overemphasis on keeping his patients comfortable above other more important aims of the treatment. His notes had an unusual amount of concern for his reactions to the patients and what he did to keep himself comfortable. (3) He unnecessarily attempted the authority-figure role. The role was carried off poorly. (4) He expressed to one patient his disappointment in what the patient was doing. One patient expressed to him much confusion about what he, the therapist, was doing to him.

*Resident No. 3 (High):*

Dr. Wright was a tall, dark-haired man of twenty-eight years who had been married seven years and had three children. Six out of thirteen supervisors emphasized his "good judgment," his "good relationships with patients"; five his "knowledge of psychiatry and related fields"; four his "improvement" during their supervision; four his "objectivity"; four his "honesty"; four his "quietness"; four how much they liked him; four his "intelligence" (although his IQ was actually lower than that of any other High and most of the Lows).

Again, supervisors' statements that are difficult to categorize reveal the man most distinctly. In his first year of residency, one of his supervisors said, "In spite of his shy, quiet manner I know he is a bright and sensitive fellow. In addition to this he has a sound medical background and good judgment in using it. My chief criticism of Dr. Wright perhaps is that he is too inclined to be passive and not to assert his real feelings, particularly if they clash with those in authority. But I think this will improve in time." A training analyst who functioned as a supervisor said of him then, "The first impression is that of great reliability, seriousness and deliberateness in his judgment; he wouldn't come out quickly with a decision or an opinion, but one has the feeling that it is very well thought over or through—more intellectual than intuitive but very solid. He strikes me as one of the best of those I've seen. I think he has the right personality to convey friendly paternal authority, firm but friendly. Certainly he has an excellent grasp of psychiatric problems. I like him very much."

In the last year of his residency a supervisor commented: "He is becoming much more open, much more verbal in communicating his ideas with me and with others and I find him to be an intelligent and mature person who is not disturbed by emotional relationships with patients." Another supervisor who said in the first year, "He is a very good man with a peculiar personality quality that leaves him just short of definitive decision and action in dealing with some patient situations," said in his later evaluation, "He seems to be maturing every day and I think he is a very sound man."

*Resident No. 4 (High):*

Dr. Haspel was a slight, twenty-five-year-old married man. Five of his twelve supervisors emphasized his "conscientiousness"; four his "intelligence"; four his "great sensitivity and empathy"; three his "warmth"; three his "honesty"; three his "potentialities," and three his "passivity."

In his first year he was described by one supervisor as "professionally mature which is in sharp contrast to the general impression he makes physically of immaturity. He has excellent clinical judgment and establishes satisfactory relationships with all his patients and their relatives." Another supervisor says, "He is serious and conscientious; the patients have confidence in him; he is responsible. He can take an awful lot of aggressiveness from the alcoholics sometimes, as well as a lot of passive dependency." In his last year of residency one supervisor said of him, "He easily becomes aware of problems in himself and is able to solve them; although he is very serious he always can find humor in situations, and seems to enjoy dealing with patients a little more than most of the residents. For example, when a patient chastises him for having given the wrong advice when he has made no such comment, he understands the patient in the situation and while away from the situation he sees the humor in an adult's using a child-like mechanism." (See also Dr. Abbott, another High, discussed in Chap. 11.)

#### COMPARISON OF FINDINGS FROM EXPERIENCE SURVEY AND EMPIRICAL STUDIES

Let us see, finally, how well our various sources of empirical evidence agree with each other and with the consensus of experts gathered in Chap. 15.

First, there is a sizable group of variables on which general agreement exists. It is universally found that a high order of intelligence is needed for psychiatric work; opinion, ratings, observations, and measurement by intelligence tests all converge on this point. The closely related concept, clarity of thought, was not mentioned by the experts except as a quality of a good psychoanalyst, but both our correlational results (using pooled personality ratings) and supervisors' observations specified this quality of intelligence. Empathy may be defined in many ways, but everyone agrees that it is necessary for the psychiatrist. We did not find that integrity played as prominent a role in distinguishing among residents of varying levels of competence as might have been expected from the unanimity of our

experience survey, perhaps because of preliminary screening. The capacity for good relationships with co-workers was confirmed all down the line, though much less was said directly on this point in the opinions on necessary qualities for psychotherapists and psychoanalysts, who so often work alone. And the readiness or capacity for growth and change is the last quality on which our two sources of empirical evidence and four types of expert opinion \* all agree.

Very substantial agreement exists on another group of qualities. Though our correlations suggest that verbal capacity has only a modest amount of pertinence to psychiatric performance, it was mentioned by supervisors and by all categories of experts except those who discussed the finished psychiatrist. It is curious that all sources of evidence *except* opinions on qualities to be sought in residents agree on interest in psychiatry and objectivity in relationships to other persons, particularly patients and supervisors (and their general models—dependent figures and authority figures). It is understandable that all sources of evidence except opinions on requisites for psychotherapists and psychoanalysts should agree on likeability, again because of the frequently isolated nature of therapeutic private practice. It is a little more surprising that opinions on qualities to look for in psychiatric applicants or in psychiatrists themselves should not mention self-objectivity and spontaneity.

The fifteen aspects of personality just mentioned make up the group on which the two principal empirical sources of data agreed. On each of them, a sizable body of opinion concurred. On two other groups of qualities, expert opinion agrees with one of our types of data but the other yields no positive results.

Seven such variables, recommended by expert opinion, were rated by the Predictive Study judges and correlated significantly with Over-all Competence. Of these qualities, five are not likely to show up directly in the behavior observed by supervisors: tolerance (or the lack of ethnic prejudice), consciousness of social injustice, cultural wealth, freedom from material preoccupations (being motivated by desires for money and comfort), and mature heterosexual adjustment. Psychological-mindedness was occasionally mentioned by supervisors but not frequently enough to be separately tallied; it was grouped with such terms as "understanding," "intuition," "tact," "good insight into patients" under the heading of sensitivity

\* On qualities of desirable applicants for psychiatric training, of psychiatrists, of psychotherapists, and of psychoanalysts.

and empathy, since all these concepts seemed to be used rather loosely and interchangeably. Similarly, the system set up for coding supervisors' comments did not provide explicitly for freedom from status-mindedness and other aspects of the authoritarian personality pattern, which were split between "poor relationships with patients," "poor relationships with co-workers," and "lack of kindly interest in people."

The group of qualities on which the supervisors' descriptions alone agree with expert opinion is larger, but there is just as little evidence of any disagreement. Most of them are qualities that were not included in the list of thirty-two personality variables but appeared instead among the minor predictive variables: the ability to win respect and trust, firmness and authority, emotional warmth, acceptance of responsibility, spontaneity, sensitivity, and judgment (included under our variable, Judiciousness). Maturity and emotional stability are quite general concepts, which seemed to have been adequately covered by a number of more specific ones (especially Adequacy of Emotional Control). We did not assess either imagination or initiative and energy: in the case of the latter, this seems to have been a mistake (see retrospective studies in Chap. 14). We did not distinguish among imagination, originality, and creativity but, in this whole area, decided to focus primarily on the low end of the scale, (Freedom from) Stereotypy of Thought (see below).

There were twenty-one concepts mentioned by experts in relation to two or more of the four groups (residents, psychiatrists, therapists, analysts), but no evidence about them (either positive or negative) emerged from the empirical studies reported in this chapter. (The full list is given in Appendix 16.3, Table 16.3.) For the most part, they are near-synonyms for terms on which there was good agreement (*e.g.*, psychological understanding and intuition—*cf.* Empathy and Psychological-Mindedness.) A number of them are types of psychopathology which were screened out (with only occasional exceptions) by the Admissions Committee procedures and so could not be expected to differentiate in the population of accepted residents on whom our empirical studies were based. One other emphasis deserves special mention: many of the experts in our experience survey seem to have been impressed by *breadth*—breadth of interests, breadth of life experience, broad liberal education, richness of personality are all concepts that occur again and again in Tables 15.1 and 15.2, but not in Table 16.2 (comments by super-

visors in the Menninger School of Psychiatry) and not in the Predictive Study variables. We think it likely that the emphasis on breadth is part of the desire to fill one's profession with the finest kind of people rather than being the product of observation that these are necessary qualities for the practice of psychiatry.

We now come to a few outright disagreements. Creativeness and originality (just as leadership ability, perhaps) can hardly be opposed by anyone, and the field needs a good many people who have them, but are they necessary to the average practitioner? Even though the supervisors described the residents they rated highly in these terms, the pooled ratings by Judges I and II of Freedom from Stereotypy of Thought (on which high ratings were earned by signs of originality) correlated only .15 with the criterion. It may be that psychiatry is by its very nature more creative than nut-and-bolt assembly, but the degree of imaginative originality demanded in the day-to-day work of the resident is not enough to warrant special attention to this variable.

A good sense of humor is another desirable quality, mentioned by several expert sources but not by the supervisors. The pooled assessment ratings of this variable failed to correlate better than zero with psychiatric competence, despite the fact that a number of Highs had (in our judgment) excellent senses of humor. We still think that it is desirable for a psychiatrist, as for anyone else, to have a sense of humor, but apparently it is not absolutely necessary.

The apparent disagreement on sensitivity is quickly resolved; the assessment ratings were out of step, and the difficulty clearly was with the definition of our variable, Perceptual Sensitivity, which emphasized a general alertness to perceptual detail. Some kind of sensitivity is obviously desirable and probably necessary for a psychiatrist; we have at least shown that "observational ability" is not so generalized that all test behavior that might be so classified is relevant.

The major disagreements come in the important but elusive dimension of motivation for psychiatric work. Curiosity is the motive most emphasized by the experts from Freud on down, whereas the need to help tends to get less unequivocal endorsement from experts, some mainly warning against it in its excessive forms. Yet, in the correlations the Need to Help is one of the best correlating variables and Psychological Curiosity one of the worst! The apparent disagreement between predictive assessments and super-

visors' observations (their descriptions of "kindly interest in people" did not differentiate Highs and Lows) may be due to the fact that supervisors' comments on overstrong desire to help, to coddle patients, etc., were coded along with the more appropriate expressions.

Curiosity is the real surprise. Our pooled ratings of it not only completely failed to give any significant validity but three of the four correlations for the two judges were slightly negative—whatever tendency there was may have run in the wrong direction. What is the difficulty here? Is it that this motive is really irrelevant? We find it easier to believe that we may have done a bad job of assessing it. The diagnostic testing experience of our judges had oriented them toward picking up test signs of voyeurism, and they may have been unduly influenced by such signs, paying too little attention to the degree to which the curiosity indicated was sexualized.

Another possibility is that a degree of curiosity out of the ordinary is not needed for ordinary psychiatric practice and that it has been stressed so much by writers because they themselves had an unusual interest in research and, thus, probably strong curiosity.

There seems to be disagreement on six other qualities, all of them more or less pathological contraindicators that did not distinguish better from poorer residents in our correlational or observational studies: fear of one's own aggression, schizoid trends, projective and paranoid trends, rigidity, and anxiety (too much or too little). The disagreement is probably not real for two reasons: first, the empirical studies dealt with a group of acceptees who had already been screened by tests and interviews which, generally, caught serious pathology; second, these are all quantitative matters, and it may well be that each of them is a valid contraindication when extreme enough. It is worthy of note, however, that any of these pathological features may be apparent at the time a man applies for training without being sufficient reason to disqualify him.

The final disagreement is familiar to us from the discussions of interviewers' errors in Chap. 14: general appearance and manner. Quite a number of authorities believe that this kind of thing is important; we feel confident that if they subjected their judgments to rigorous validation from independent follow-up evaluations, they would find that prepossessingness or its lack is mainly a source of error. Parenthetically, a word might be said about aggression and hostility. Expert opinion does not condemn it, and our evidence is that good residents as well as poor ones may display a fair amount

of hostility, yet the studies of erroneous prediction in Chap. 14 showed that occasionally interviewers formed mistakenly negative impressions of aggressive applicants. A particularly noteworthy example is one of the Highs, a man who did extremely well as a resident and since then as a psychiatrist: he was very nearly rejected because of his irritability and hostility during assessment, which was noted by psychiatrists and psychologists alike. It happened, however, that he was under strain because of a family crisis and there had been a misunderstanding about his application so that he thought he was being offered a staff job and did not expect to be interviewed and tested. All of this came out in a final interview, fortunately, and the Admissions Committee correctly trusted the judgment of the interviewer who urged that the negative impressions of all the others be ignored.

Finally, a few qualities turned up in this project which had not been mentioned as such by authorities but which showed a relation to the criterion. Mature Heterosexual Adjustment, as rated by Judges I and II, correlated well with Over-all Competence; it is only tangentially referred to in recommendations that overt perversion should be a contraindication or in occasional references to a happy marriage as a good prognostic sign. No doubt it has been neglected because there are many exceptions, particularly certain unmarried women who have been outstanding psychoanalysts. We were dealing with an all male population, however; even if there are a number of exceptions, the same is true of almost all of the variables we have been considering: all of the correlations were relatively low, even when significantly different from chance. Attitudes and orientations toward women and love relationships deserve careful assessment.

*Emotional Appropriateness* is a term we used to cover such matters as tact, certain aspects of sensitivity, and kindred concepts. This way of conceptualizing these issues has, perhaps, the advantage of focusing on the degree to which impulse and affect generally are modulated, adaptively attuned to the demands of situations. On the side of response, it is the counterpart to empathic receptiveness, which it presupposes.

We have discussed above the somewhat novel conception of warmth that comes from our supervisors' observations. It is backed up by two others of the latters' differentiating categories which had not been directly mentioned by authorities we read or consulted:

self-possession and inhibition (though it is true that the last term did not discriminate to a very significant degree). Under these headings were collected observations that the better residents were calm, self-controlled, and composed; their equanimity might often cover a fair amount of inner conflict, and the effort at control may have produced more inhibition than might be expected. The tendency for the Highs to be described as passive may also be part of this pattern of quiet and reserve. Again, of course, this is not to be taken as a universal desideratum, but it has apparently received less emphasis than it deserves.

To sum up: evidence from our Predictive Study and from supervisors' observations for the most part confirms recommendations made by authorities from their experience in selecting and training psychiatrists, psychotherapists, and psychoanalysts. It serves as a corrective to certain emphases that seem to be mistaken and has turned up a few ideas about the personality of the desirable psychiatric resident that are worthy of trial and further exploration.

In the above analysis, it has been necessary to deal with personality in terms of isolated traits, considering the predictive influence of each by itself. This may give rise to the impression that the long process of personality assessment might be cut short by concentrating on the measurement of only a few of the most "valid" traits. We very much doubt, however, that anyone could achieve good measures of even a small group of traits without a thorough clinical assessment of each man.



P A R T I V

CONCLUSIONS  
AND  
RECOMMENDATIONS



## RECOMMENDATIONS

DURING THE YEARS FROM 1947 THROUGH 1957, WE DEVOTED A GREAT deal of time to thinking and reading about the selection of psychiatric residents in addition to carrying out our research. Since 1952, each of the authors has had the opportunity to serve for one or more years as a regular member of an Admissions Committee of the Menninger School of Psychiatry, grappling with day-to-day practical problems as well as continuing to analyze the research data. Consequently, we have built up a store of knowledge, opinion, and (doubtless) prejudice on the subject. Faced now with the task of making specific recommendations, we feel impelled to go beyond what we can back up with attested research findings and to draw upon all our experience, direct and vicarious, in setting forth the practical fruit of our work. It would be much easier to stick closely to statistically reliable findings. We are aware, however, that residents must and will be selected somehow and that those who are doing this work would like to have any assistance that is available. With a *caveat lector*, here it is.

Before making specific suggestions, let us review some of the pressures on a committee selecting residents for a training program in psychiatry. First, there are the inseparably intertwined demands of time and money. Even if more extensive and conclusive research than ours should determine that it pays to spend several days in assessing each applicant, many institutions would be unable to do so.

The most skillful clinicians in any group are usually in greatest demand for service and training, so that it is difficult to keep them on an Admissions Committee, much less to resist the demands to cut down the amount of time that they devote to it. Inevitably, such men are also the most highly paid, and if they devote many hours to selecting residents, this forces up the cost of the selection procedure.

There is consequently a pressure, from the nature of an institutional setting, toward cutting down the amount of time spent in assessing any one candidate, as well as a tendency for other duties to pre-empt the time of the most skilled clinicians. A regard for the situation of the candidate exerts a similar pressure against extensive assessment procedures. At a time in his career when typically he has neither time nor money, he is asked to spend a good deal of both to be evaluated.

At the same time, the work of an Admissions Committee comes under the influence of another, somewhat antithetical, set of forces. It is no trivial decision that is being made when a committee decides whether or not a man may enter a professional school. The procedures used have to have a degree of "face validity," therefore, or the impression will spread that a serious task is not being handled right. It is quite conceivable, for example, that in a particular institution a good job of selecting residents could be done merely by a careful evaluation of credentials and self-administered psychological tests. Many psychiatric admissions committees would be extremely loath to follow such a procedure, however; they would feel that they were taking people blindly and not giving them a fair chance. Applicants would not only fail to respect such a procedure but would not get a chance to talk to some of the people who represent the institution and work out the inner process of decision. For these reasons it seems likely to us that the psychiatric interview will be retained as a selection procedure whenever possible. Our recommendations take this for granted and are addressed, in part, to the task of making the most out of the interview.

#### GENERAL RECOMMENDATIONS ON SELECTING RESIDENTS

##### INTERVIEWING

Everyone has his own style of interviewing; constraining it by a fixed schedule of questions is not advantageous. A general outline

in which several important areas are covered has proved useful. The aim should be to get the applicant to talk freely within these areas; for example, about what kind of person he is and the persons who have been important to him. The interviewer should weigh his findings in terms of the personality qualities we will describe below.

Our experience, particularly with predictive failures, makes us quite skeptical of the value of "stress interviews" or similar procedures, the essence of which is to put a man on a very unpleasant spot and try to learn about his "real personality" from the way he squirms. When a man is to be considered for training in psychiatry or psychoanalysis, we need to predict future behavior from a small sample of present behavior and reports of past behavior. We can do so by inferring some inner structures and deducing the behavior that they would lead to in other situations, or by simply extrapolating the trends we see to the future, or by both. If present behavior is taken as a sample of what is to come, a representative situation should be provided. A very large part of a psychiatrist's or analyst's behavior will take place in situations that offer little stress, or a very different kind from that produced by the evaluative eye of an Admissions Committee.

The above is not just a logical argument that one cannot predict active behavior from passive reactions to stress. Our findings also indicate that many men who seemed, when assessed, to be too lacking in self-confidence, or too brassily aggressive, or too intellectually pretentious in this threatening situation turned out to be quite different in the actual business of confronting patients.

A related error is to concentrate too exclusively on piercing defenses to explore the conflicts and pathological trends hidden underneath. Very often it happened that testers and interviewers alike underestimated candidates because they succeeded so well in seeing through defenses to underlying weaknesses, failing to gauge how effectively the ego's adaptive tendencies operated in real life. This is not to argue that one is better off not knowing what goes on under the superficialities of personality; only that one should not underestimate what can be done with a good set of defenses and that when considering deeper material he should do so carefully in light of what the person's resources are for everyday functioning.

At the same time, the studies of interviewers' errors reported in Chap. 14 indicate that many interviewers may be too much influ-

enced by apparent warmth, self-confidence, deep clear voices and commanding presence, or other such superficial aspects.\* If the kind of psychiatrist being trained is primarily a psychoanalytically-oriented psychotherapist, the interviewer should try to ignore those features of appearance and manner that constitute attractiveness and concentrate on getting at the intrinsic qualities of the personality. If, on the other hand, one is looking for someone to be an institutional leader, who will spend more of his time in administration and influencing colleagues, the requirements may be somewhat different. Under such circumstances, it probably helps to get a man who can make an impressive appearance, who can speak well and be sure of himself in brief contacts with many kinds of people.

Regarding the actual conduct of the interview session itself, there are many styles of effective interviewing, and it is best for an interviewer to develop a technique with which he is personally comfortable. Some may find it helpful, however, to hear how Lionel Blitzsten, a training analyst with many years of experience, described his own technique (in an unpublished letter to Robert P. Knight):

When an applicant first enters my office, my first attempt is to put him at ease by offering him one of two comfortable chairs—a pair—in the other of which I sit. I then ask him if he would care to smoke. I pay strict attention to his entire manner of entering the room, his opening comments, whether they are about the office, objects in it, or anything specific; in other words, I observe him as I would anyone who comes for a psychiatric interview. I then proceed to tell him the real purpose of the interview, assuring him that this is not a "third degree" and that he is not "on the spot." I do, however, tell him that the interview is primarily in his own interest and only secondarily in the interest of psychoanalysis, but that it is in the nature of a psychiatric interview which will enable us to get some idea as to whether he should subject himself to a personal analysis—or if we think he should; to get some idea of the difficulties, the length of time it might require, and to what kind of an analyst he should go.

I then proceed as I would with a patient and tell him to tell me something of himself as a person (leaving a recital of his formal training to the last; they usually begin with this). Of course, from what he says I take my cue as to what topic I ask him to elaborate and I usually direct him to talk about the important symbols—parents, siblings, and their surrogates—to get an idea of his re-

\* Especially if their attention is directed to such irrelevancies by a report form.

sponse to discipline, authority, competition, rivalry, his own goals and ambitions.

A second very important item is to discuss their life commitments which might have lasting consequences, and their own evaluation of them . . . more specifically, marked deception, mishandling of money, marked antagonisms and broken relationships, type of overt perversion if current, and with whom, etc.

Third, a discussion of their cultural background—hobbies, avocations, other interests aside from their profession, to get an idea of their capacity to objectivize instinctual impulses present in all of us and the savor of the integrating power of the ego. In over 100 recent applicants I have been appalled at the sterility of their lives—the almost complete absence of any cultural background and any marked interest in the current socio-economic scene. I think that a cultural background—not in the narrow sense—is absolutely necessary for a psychiatrist, and helps empathy and psychological understanding for it enables him to identify himself with the other individual's life and, ergo, to have some feelings for his needs or frustrations within the scope of his social framework. I am not a perfectionist about this. I would far rather accept an applicant who has some cultural interests and artistic ones who has had severe neurotic or even near-psychotic disturbances than one who is lacking in this and whose life is a model of so-called adjustment. I agree entirely with Hanns Sachs that one criterion for *not* accepting an applicant is that he be a so-called "normal person" who has no conflicts, who gets along beautifully with everyone, etc. These are frequently persons with such a rigid façade—a necessary one—that to penetrate it would lead to disintegration. I have found some of these to be potentially dangerous paranoids.

Fourth, I usually ask every candidate how he would feel if he were not accepted. The reaction to this question sometimes gives more valuable clues to the character of the individual than anything else, especially to those just mentioned above.

Fifth, I always have every one discuss his economic situation, especially with reference to the realities of a long term apprenticeship to psychoanalysis, in relation to their own social living and that of their families. Here one gets the best ideas of the maturity of the individual and a real testing of his capacity and ways of meeting real situations.

Before they leave I frequently ask them to speak briefly about their feelings about the interview. This is a valuable item.

Such an interview requires a minimum of an hour, and frequently, if the applicant is a person of some prospect, I take more time—as much as two hours or more, even if it requires further interviews. This to me is an obligation of any member of a training committee if he is sincere about his task.

**USE OF THE TEST BATTERY**

Our results strongly suggest that, within limits, the nature of the particular tests used matters less than who uses them. But it is not enough simply to get a "good tester"; he may not be any good at this unique task. With interviewers and with psychological test interpreters alike, *one cannot know who can select psychiatrists until they have been tried out at the job itself.*

On the other hand, it is certainly possible to underestimate the importance of the particular tests used. People who are not actively engaged in diagnostic testing naturally enough slip into the more or less involuntary assumption that a test is a test—or, frequently, that psychological testing means giving the Rorschach and maybe something else. To consider comparable the results based on two different sets of tests would be as inappropriate as it would be to conclude that the interview had proved worthless if an interviewer who read off a list of questions, taking down "yes" or "no" answers, and then recording his judgments, failed in his predictions. (Let us hasten to add that this is a purely hypothetical case!)

It is often remarked these days that the instruments of the diagnostic tester do not deserve the name "tests." A test of spelling gives you a quantitative score, which—if it is a good test—is a direct indication of spelling ability. By contrast, a projective technique is much more like an interview. It is a more or less standardized situation to which a person can respond under the eyes of a trained administrator who is able to note in the behavior, verbalizations, and test responses a rich and revealing segment of human personality in action. Psychological testing of the kind used in the present project falls somewhere between the objectivity of the unidimensional, quantitative spelling test and the subjectivity of the clinical interview. A great deal depends on the sensitivity and talent of the tester, yet no matter how skilled and experienced he may be, he has to have adequate materials to work with if we expect him to provide us with useful answers to clinical questions.

It is advisable to give *several* tests and ones with which the psychologist is familiar and competent. Projective and intellective tests, evaluated along with a careful study of credentials, should give the best validities. If, for example, one could give only the Rorschach and TAT, it would be better to give no tests at all rather than spend time with so dubious a prospect of satisfactory results. Projective tests

give valuable insights into personality, but the level of material from which they draw varies so much from one case to another and its significance is so dependent on a framework of realistic knowledge about the person that projective techniques can make their proper contribution only when used in conjunction with other methods, preferably including tests of intelligence and association, and historical material.

The logical implication of some of our findings is that psychological testing will be most useful when carried out by clinicians who have full access to a maximum of non-test data on the applicants. Blind testing, such as we had to resort to in our Predictive Study, may be an impressive stunt, a useful learning device, or a stratagem to convince those who are disinclined to believe that the tests themselves add anything to clinical impressions. A mature and serious approach to selection, in which the only goal is to make best use of time and information, requires that the psychologist interpret his test results along with a careful study of credentials and any other such information as may be conveniently available. The result will be not a pure measure of how much comes from the tests themselves but maximal selection efficiency. Some psychologists find that they can follow "the inherent logic of the test data" (Schafer) better if they work with a minimum of extra-test information, bringing in the rest of the clinical data only at the end. We see no objection to such a working procedure so long as the synthesis is ultimately made.

#### QUALITIES TO LOOK FOR IN ASSESSING APPLICANTS

There is no one type of person who makes a good resident, nor is there any single type to be avoided. Rather, there are many different kinds of people who can be useful in psychiatry in one of its numerous branches. Most of what follows, however, is directed toward *selecting residents for training in psychoanalytically oriented psychiatry with a heavy emphasis on psychotherapy.*

A special word is perhaps in order on the long-run objective which, of course, is to find and train men who can become good psychiatrists, not just good residents. Some may be difficult to handle as students in a residency training program yet do well in practice. Men who lack certain specific abilities of leadership, or who find it difficult to adapt themselves to a subordinate role in such a complexly interdependent group situation as we find in a modern psychiatric hospital, may nevertheless do very well in individual private practice. Moreover, there are dull-looking, inhibited people who

blossom out during training and develop qualities the latent seeds of which are difficult to assess at the time of application. (See Chaps. 5, 6, and 16; also see the concluding section of Chap. 16 for a summary of evidence on the relative weights to be given the various qualities discussed in this section, below.)

**MOTIVATIONS.**—In considering this general topic, as well as other areas, we shall follow the organization given by our list of thirty-two personality variables, presented in Chap. 10.

Essentially two matters are considered under the heading of Freedom from Material Preoccupations: the amount of monetary motivation and the candidate's interest in psychiatry as a potentially comfortable form of medical practice. Carefully evaluated, both of these considerations can be useful though they should not be given heavy weight. Mild and moderate degrees of pecuniary interest were found at all levels of competence in our group of residents, but when there are enough indications of concern with money to suggest a real preoccupation, or occasional indication of interest in psychiatry primarily as a way of making "big money," the prospects are not at all good.

It is difficult to state what constitutes preoccupation or otherwise, to indicate where the critical point on a continuum of pecuniary motivation lies, without giving extensive quotations. If, in an interview, a man mentions making money with particular emphasis (such as its being one of the first things mentioned) when asked his reasons for wanting to go into psychiatry, and if his other reasons seem lacking in conviction or simply not present, there should be little doubt that his motivation is undesirable. As a rule, however, candidates are sophisticated enough to realize that they will make a better impression with more idealistic kinds of stated motives. Consequently, the indirect indications of the projective tests become particularly useful in such cases. If there is any delay, disturbance, or unusual reaction to the word "money" in the Word Association Test, it should be inquired into. In the TAT particular attention should be given to a massing of stories in which *economic* hardships of the characters are particularly stressed or the projected motivations very conspicuously include making money. In all cases, however, attention should be paid to the realistic situation.

If a man has had to work his way through medical school and is struggling financially at the time he is assessed, it is understandable

and allowable for him to betray more concern with money than his financially more fortunate peers. There is no substitute for an accumulation of experience with the kind of population one is assessing. Against such a background, the man who is really greedy for money-making opportunities will stand out clearly.

It is fairly unusual for candidates to discuss at any great length considerations of comfort and convenience when being interviewed on their motives for going into psychiatry, unless they have been specifically directed to the specialty because of a condition such as heart disease or arrested tuberculosis, which preclude the physically more demanding specialties. In our experience the few men who have made much of psychiatry as an easy form of practice have generally been inadequate residents.

**Freedom from Status-Mindedness:** Together with acquisitiveness, the craving for status represents the most common form of extrinsic motivation in our society. Clearly a person who mainly wants recognition and power has no specific vocation for psychiatric work. He is not likely, therefore, to get much gratification from the most constructive aspects of his role as a psychiatrist. Indeed, motivation of this type will tend to make it difficult for a resident to learn not to impose his own goals on a patient but to allow the patient to develop in his own way; in general, it is conducive to an authoritarian type of role in psychotherapy, in the management of patients, and in the administration of a ward or hospital. In addition, the many ways in which patients try defensively to demean the therapist are likely to provoke anxiety and retaliation in the status-oriented resident. It is quite possible to get results with some patients using an authoritarian approach, but we believe that they are less well-founded, shorter lasting, and of a different quality than the ones achieved through helping patients to assume responsibility for themselves.

The status-minded applicant is likely to show this orientation in his attitude toward the interviewer or tester in almost any situation. If he perceives the other to be his superior, he is likely to be ingratiating or submissive, or else competitive and rebellious. It is particularly noteworthy if a man, who has demonstrated respect for an older psychiatric interviewer, shows impatience and arrogance toward a clinical psychologist or social worker who may also test or interview him. Many indirect indications of status-mindedness emerge in interview discussions of motivations for psychiatry, social and political attitudes, and personal history. In the Rorschach test,

symbols of power and authority often appear in the records of status-minded applicants: crowns, coats of arms, military insignia of high rank, "a despotic king," and similarly elaborated responses indicating identification with authoritarian models. TAT stories may feature themes of "thirst for glory" and fame as the prime motivation of heroes, or themes of shame and chagrin at losing face; however, these should be treated cautiously unless they are blatant or frequently repeated.

In assessing the Need to Help it is important to distinguish between genuine helpfulness, with a strong component of love in it, and pseudo-kindliness based on guilt or reaction formation against hostility. The kind of excessive therapeutic zeal against which so many analysts, following Freud, have warned is clearly a defensive or narcissistic maneuver, a form of acting out with the patient rather than a direct response to his realistic plight. The genuine need to help is not nearly so showy a phenomenon, so it has perhaps been overlooked too often. According to our impressions, the desire to help others more easily develops into a warm and compassionate, yet appropriate, kind of motive when it is based not upon sado-masochism but upon an identification with the sufferer, where the underlying need is a passive or oral one—provided that it is well-modulated.

It is important, therefore, to distinguish carefully in the interview between pseudo-compassionate protestations and a well-modulated desire to heal. The quality of protest and exaggeration should be lacking, and the degree of stated motivation should be consistent with the reported history of other more general interests and activities. A spontaneous report of real gratification from being able to help suffering people in one's internship is a positive example. The statement that the choice of specialty fell between psychiatry and surgery is likely to be a negative one. It is good to inquire into the applicant's feelings about children—his own (or his desire to have them) and those of his friends, going into his interest in working with children or just in being with them. Again, in doubtful cases the quality of the need to help projected in the TAT stories can be helpful, particularly in relation to general test indications about the way that hostility is handled.

A closely related issue is that of emotional warmth. Judging from our study of better *vs.* poorer residents (Chap. 16) and the tendency of interviewers to overestimate candidates they considered most outstanding in warmth, interviewers should be more impressed by

indirect or quiet evidences of compassion than by a strongly manifest impression of warmth directly experienced in the interview.

It is desirable if an applicant has Psychological-Mindedness, but quite a number of people who did not give much evidence of it when they were assessed developed this quality, or at least made it manifest, in the course of their training. Definite positive indications are therefore to be given positive weight, but a mere lack of them should not be considered a black mark against an applicant unless it is reinforced by strong indications of emotional resistance against psychological explanations, a rigid personal investment in an organic orientation, or the like. What must be present in the applicant, therefore, is not necessarily fully developed psychological-mindedness but a kind of defensive structure and personality setting in which it can easily develop. Any of the following may characterize a person in whom psychological-mindedness is *not* likely to develop: an authoritarian personality structure; intense prejudices; heavy emphasis on repression and narrowing of the ego as defenses; an aversion to introspection, such as may be found in addictive or impulsive character disorders; or a value system emphasizing externals, obvious symbols of power and wealth, sports, and superficial sociability at the expense of the life of the mind and cultural interests.

In the interview it often helps in the evaluation of psychological-mindedness to get the candidate to describe and discuss a patient, members of his family, or some person known to both him and the interviewer. Sometimes organic-mindedness can be discovered by asking rather direct questions about interests in neurology, somatic treatment, and the like. Of course, these are legitimate aspects of psychiatry, and the question to be decided is whether they *exclude* an interest in the more central, psychological aspect of psychiatry. Another approach often used is to discuss the applicant's taste in literature to see whether his interests have been directed toward fiction and biography in which character delineation has been particularly important—but only if he mentions these features and can back up his evaluation by concrete discussion.

Psychological testing offers particularly rich opportunities to observe both positive and negative indications of psychological-mindedness. The inquiry to both the Word Association and the Rorschach tests throws valuable light on the subject's capacity for introspecting on his own thought processes; and the latter test gives, in addition, formal indications through the ability, or lack of it, to describe the

use of shading as a determinant of responses. In the TAT psychological-mindedness shows up in the adequacy with which characters are developed, instead of appearing as simple cardboard figures, and particularly in the subtlety and complexity of motivation and interpersonal relationships in the stories. But the Self-Interpretation is par excellence the procedure that can give clear indications of developed psychological-mindedness, or of strong defenses against it. The best single cue to psychological-mindedness in the Self-Interpretation is the subject's spontaneous acknowledgment that he suppressed certain material or otherwise changed a story while telling or writing it. Even the Self-Interpretation, however, has not in our experience been able to detect latent psychological-mindedness in a person who displays no conspicuous signs of it either positively or negatively.

Psychological curiosity may not be quite as important a requisite for psychiatric residents as has often been stated (see discussion of curiosity in Chap. 16), and, similar to the need to help, it can be satisfied directly in the processes of diagnosis and therapy. If strong interest in psychiatry is present, however, it is probably superfluous to demand that the applicant be able to verbalize a felt need to find out about people and their behavior.

From comments by supervisors we learned that clearly sexualized curiosity was not infrequently found in inadequate residents, whereas no such comments came up in the discussion of superior ones. Curiosity, again like helpfulness, can become intimately bound up with defensive, security-oriented operations or with acting out in the service of relatively direct instinctual gratification, either of which is likely to be harmful to the patient and not in the best interests of good professional work. Therefore, we suggest that the usual clinical indications of relatively unsublimated voyeurism be considered *negative* indications of fitness for psychiatry.

Perhaps the most positive indication of the useful kinds of curiosity is an expressed interest in research. Good residents more often than poor ones indicated a desire to spend part of their time in research work during or after training, and they were more likely to have done some type of investigative work in the past. In view of the small supply of research psychiatrists and the particularly great need for them, special consideration should be given to applicants who are interested in research in more than a vague or superficial way, especially those who have done some research in whatever field.

A question about past research activities can easily be included in an application form (as was done in the Menninger School of Psychiatry); otherwise, information on this point usually has to be obtained through interviewing.

In our original list, the variable Internalized Interest in Work tended in use to be fused with Internalized Interest in Psychiatry. We might have done better to have substituted "achievement motivation" for the former. Reports from teachers and supervisors in letters of recommendation can be useful here.

It is advisable to inquire into the kinds of jobs the applicant has held and how he performed in them. Indirect, but nonetheless valuable, indications come from the psychological tests, particularly an intelligence test such as the Wechsler-Bellevue. Here in microcosm one can see how a potential resident attacks an assigned task, with what perseverance he carries through in the face of difficulty. Some-what similar observations may be made in the projective tests, but the TAT is particularly useful for projected indications of genuine achievement motivation. A person with a positive orientation toward achievement characteristically tells stories in which the hero not only attains goals but does so by means of realistic, adaptive efforts. If there is a quick and easy jump in the story between a stated goal (such as becoming a great violinist in card 1, or a surgeon in card 8BM) to the full realization of the goal without any emphasis on the intervening years of preparation and work, the achievement motivation cannot be said to be strong, and the subject's work habits are likely to be poor. On this last point, however, work habits should not be the focus of much inquiry. It is, of course, a good thing if a man applies himself to his work conscientiously, but this quality did not distinguish better and poorer candidates and may lead the Ad-missions Committee to overestimate poor applicants.

Regarding the variable Internalized Interest in Psychiatry, we made in part a rating of the relative strength of the positive, intrinsic motivations just discussed as against the negative, extrinsic ones. But in the Interview, and to some extent in the Credentials, it was possible to find more direct evidences of expressed interest in the subject matter of psychiatry itself, and it seems worth while to emphasize the desirability of paying attention to this in interviewing. In Chap. 1, we saw that a survey of major training centers showed none in which the candidate's expressions of interest in psychiatry were considered to be a matter of principal concern in assessing him, yet this proved

to be a very useful variable in our own work predicting several criteria at all levels of competence. It may be that it was not emphasized in the survey because it may have seemed almost too obvious, but if a man's interest is not intense and sustaining, he may leave the field at some point of discouragement, boredom, or stress.

From credentials and the interview we can learn whether or not the applicant's interest in psychiatry has been of long standing; whether the impetus toward this field came from the natural development of his own interests or from prodding by parents, teachers, or peers; whether or not he has a general pattern of interests in psychoanalysis, psychology, and the social sciences and humanities more broadly. Such positive indications are important if they are present, but their failure to show up does not constitute very strong negative evidence. The simple failure to develop an interest in psychiatry until just before application is not to be considered a contraindication, but the defensive lack of psychiatric interest and its involvement in conflict may be so. Ordinarily, existing psychological tests have little to contribute to the assessment of this particular variable.

**SELF-ORIENTATION.**—The assessment of Self-Confidence proved to be a difficult task. Some of our worst mistakes in prediction were based partly on giving too much positive weight to apparent indications of self-confidence, partly on giving too much weight to indications that an applicant was doubtful about his own adequacy. We still believe that some kind of underlying, "basic" self-confidence is probably necessary, at least in a moderate degree, if a man is to make a good psychiatrist; a good *prima facie* case can be made for it. But the stressful situation of assessment, when a great deal is at stake, complicates the task of assessment by arousing a good deal of anxiety and producing (perhaps superficial) expressions of self-doubt or shakiness of self-confidence in men who later turn out to do very good work. At the same time, others who appear only adequately secure in their self-esteem may turn out to be brash, narcissistic, or bumptious.

On the whole, self-confidence is perhaps just as well ignored in assessment, except where strikingly absent.

**Self-Objectivity** is well worth assessing, though several of the men who later became excellent psychiatrists received low ratings on this quality at the time of application, and this was one of the variables on which predictive judges agreed least well. Perhaps too much

was expected of the applicants; the important thing is to look for general indications of readiness to *develop* good self-knowledge. Such indications are closely related to psychological-mindedness, general honesty, openness, and lack of defensiveness. The task of assessment is further complicated by the tendency (by no means exclusive to unsatisfactory applicants but most frequent among them) to have picked up a smattering of psychiatric jargon in which to discuss one's problems in an entirely intellectual way, without emotional involvement. Moreover, satisfactory men are as likely as unsatisfactory ones to justify themselves under probing in the interview with a tone of moral righteousness, which therefore does not constitute a trustworthy indication of poor self-objectivity.

A few suggestions about valid clues can be given, however. Blocking, evasiveness, marked defensiveness, and other signs of acute conflict, or manifest self-deception when the candidate is asked to describe himself or his childhood, or to discuss similarities and differences between himself and his parents are negative indicators in the interview. Narcissistic egocentricity—one form of low self-objectivity—can be detected in numerous ways in tests (see Schafer, 1948), particularly by self-references in Word Association, Rorschach, or TAT. In the Self-Interpretation of the TAT, concentrating on peripheral, nonpersonal determinants of the stories (such as the attempt to attribute events or characterizations to the picture or to recent impersonal and external experiences), or a direct denial that they have any personal meaningfulness, have similar value. Another unfavorable indication is the failure to take up the most important problems adumbrated by the stories. The Self-Interpretation affords positive indications of a similar kind: clear acceptance of the stories as personally revealing, and a good coverage of the important issues touched on in them, even if the candidate is not able to give completely convincing interpretations.

In evaluating Sense of Humor, the main problem seems to be one of taste. A remark during an interview or test by the candidate may seem appropriately amusing to one person, may fall flat for another, and to a third may appear to have been an unwarranted defensive intrusion into a serious situation. Certainly, being considered for a psychiatric residency is a serious situation, and a jocular or facetious attitude is out of place. One should attempt, therefore, to distinguish carefully between the various ways in which humor is used, as well as noting its presence or absence. Ordinarily, the lack

of evidence on this point can hardly be held against the candidate; he may simply be holding a witty tongue in leash, and as we saw in Chap. 16, there is no evidence that this is a necessary quality. Perhaps the most clearly appropriate occasion for the display of a good sense of humor is in the TAT (though it also appears at times in Rorschach responses). The records of superior residents were frequently enlivened by a humorous twist in one or two stories out of the twelve, or an occasional witty turn of phrase. (One of them remarked of the figures on Rorschach card III that both a breast and a penis could be seen, adding "You might call them ambisexual.") Used sparingly in this way, a sense of humor seems to afford a healthy safety valve for the discharge of tension; it may also indicate a capacity for regression in the service of the ego. A good touchstone to keep in mind in evaluating a sense of humor is whether its expressions facilitate the principal ongoing business or whether they interrupt it.

Since the Capacity for Personality Growth is widely agreed to be a crucial variable in assessing a potential resident, and one which is particularly difficult to measure, our experiences in attempting to assess it may be of particular interest.

A useful approach is to make a careful study of evidences that the candidate has been growing in significant ways in his own recent past. The main sources of information on this score are credentials (including transcripts and letters of recommendation particularly), the interview, and the Self-Interpretation of the TAT. For example, a history of poor marks in undergraduate work, with an improving record throughout medical school, suggests the gradual overcoming of internal obstacles to intellectual work; such a lead should, of course, be followed up in the interview. In the discussion of many different kinds of topics it is possible for the subject to show that he has been in a condition of essential stasis since late adolescence or that he has been moving in either a positive or negative direction. Consider the matter of relationships to parents: it is normal and expected for a young man to be quite dependent during childhood and early adolescence, to struggle to win emotional independence sometime in adolescence or soon thereafter, and ultimately to free himself from too close ties to his family. Statements that this kind of process has been going on are not always to be taken entirely at face value, however; a young man may feel that he has achieved independence merely by virtue of the fact that he is no longer living with his family, whereas projective tests may indicate that he persists in

an infantile type of dependent reliance on parental figures, with a rebellious rather than truly free autonomy. This is one example of the way in which the capacity for personality growth can best be evaluated by having access to a variety of data.

One significant way in which the lack of this capacity can be seen is evidence that the applicant has been repeating the same mistakes, continually getting himself into the same kind of scrapes, or otherwise failing to learn from experience. A discussion of his relationships with women may show, for example, a repeated pattern of sado-masochistic relationships without real gratification or one affair after another ended by a fear of growing dependence. Such fixed self-defeating trends are particularly unfavorable signs when the applicant shows no insight into them.

The Self-Interpretation of the TAT clearly presents an opportunity for a history of progressive growth to be brought out, or for its opposite to become manifest. This may take the form of statements that a reported insight was recently gained or that some aspect of personality described is a development, a relatively new feature; valid indications of this kind were particularly likely to have to do with relations to parents. The applicant who is growing and changing is likely to describe himself in *process* terms as having problems, trying to overcome difficulties, or the like rather than to use *static* descriptions that imply a fixed self-concept. Less obvious is the fact that the unconscious feeling of being caught by repetitive neurotic mechanisms can show itself indirectly and symbolically in TAT stories—especially those told to the blank card, No. 16. An atmosphere of dullness, tedium or being in an unproductive rut is a negative indication, whereas if the projected themes are occasionally hopeful or appetitive, the settings described as gay and colorful, and the action zestful and gratifying to the heroes, it is a positive sign of capacity for personality growth.\* But an unrelievedly cheery response to the TAT cards, which do have a pull for the less happy side of life, is likely to indicate too much denial.

**ORIENTATION TOWARD OTHERS.**—The capacity to be objective in interpersonal relationships is plainly a desirable one for a psychiatrist. We considered it under two aspects, Objectivity toward Authority Figures and Objectivity toward Dependent Figures. First, we looked

\* This finding suggests the possibility that capacity for growth is related to what Erikson (1952) describes as "Basic Trust."

for indications of attitudes toward authority figures as being particularly relevant to the resident's capacity to establish appropriate relations with his supervisors so that he could learn from them and function adequately in an institutional setting.

To some extent this can be gauged very directly from statements by former teachers and supervisors in letters of recommendation, and by discussions with the applicant in the interview of how he got along with such persons. More indirect evidence comes from drawing him out in interview about his parents, but in our experience this is extremely difficult to do in a way that clearly distinguishes adequate from inadequate residents. Cues relating to parental attitudes were the most likely to lose validity on cross-validation in all of the manuals in which they appeared. About the only usable indications remaining were (on the positive side) material indicating feelings of being able to compete with father figures successfully, in the TAT and Self-Interpretation, and (negative indication) a basic lack of any respect for a parent, coming out in the interview.

Why is it such a tricky task to assess attitudes and orientations that will interfere with a resident's relationships with older patients and with his institutional superordinates? We believe that it is related to the late adolescence that seems to characterize many psychiatric residents. In their middle twenties, the most promising men for psychiatric careers are still wrestling actively with what are often considered adolescent problems such as reawakened oedipal conflict, struggling to break away from emotional dependence on their parents. As a result, in the interview they may block or get upset when discussing parents, justify them unnecessarily, or attack them with great bitterness—all without the implication that they will make inadequate residents. Close ties of affection to the mother, or guilty feelings about her, or violent rejection—none of these is either a good or a bad sign. The relation to the father is easier to assess meaningfully: continued dependence, over-rebelliousness, or inability to rebel are usually bad signs. In the TAT denial of all conflict and a pervasive tone of sweetness and light when relations between sons and parents enter the stories should be viewed with suspicion. Though they are rare, genuine expressions of love for credibly described human beings, with both good and weak sides, may be considered positive indications in the interview. On the whole, however, the accurate assessment of objectivity toward authority figures is unusually difficult if one does not have evidence from several different kinds of sources.

Objectivity toward dependent figures, or patient-surrogates, is also difficult to gauge accurately. Children, members of groups with lower social status than the candidate himself, and certain animals are the main symbolic equivalents of patients, attitudes toward whom can be inquired into for more indirect measures than are given by direct inquiry. All potential residents have had experiences with patients, however, and often a good deal can be learned from asking questions about types of patients the candidate has found it particularly gratifying to work with, types of patients he dislikes to treat, and particularly his reactions to hostility from patients. Strong, emotionalized attitudes and destructive (authoritarian, punitive, moralistic) reactions to patients elicited in this way are bad signs for the development of objectivity in psychiatric dealings with them, particularly if backed up by genetic facts or data from the projective tests that indicate ways in which such trends serve a neurotic need or fit into an ingrained personality pattern. The mere lack of evidence that an applicant has developed objective attitudes toward patients and their surrogates is not too serious a drawback, however, since a good deal of psychiatric training is devoted to developing just this capacity.

Usually, psychiatrists attempt to gauge Empathy by noting the extent to which the candidate seems to be able to understand and discuss intelligently the inner feelings of people he describes to the interviewer. There are some dangers in this procedure, however, for it assumes that a person who is able to achieve empathic understanding of someone else will also be able to impart it, and this assumption is a shaky one at best. Empathy may be a kind of understanding that doesn't have to involve words, the reception of nonverbal communications from another person. But if the interviewer has sufficient understanding of himself and what he is communicating in the situation and is sensitive enough to pick up subtle indications of the extent to which the candidate grasps them, then he may be able to use the interview situation itself in assessing this quality on a non-verbal level.

On a much simpler level, gross deficiencies of empathy can be seen wherever there is communication between two persons. If the candidate talks in such a way as to indicate that he is paying little attention to whether or not he can be understood, if he fails to grasp any but the most direct indications that the interviewer does not consider a topic germane, or that he for any reason wants to get on to

something else, the process of the interview itself may reveal a lack of empathy. This type of failure in communication may show up similarly in almost any psychological test, but perhaps most clearly in the Rorschach inquiry. The failure really to listen to what the examiner says, answering questions besides the point, and repeatedly having to be told to slow down the tempo of his speech so that the examiner can get down a verbatim record are reliable test indications of the lack of at least one type of empathy.

We are less certain about formal and projective signs that are supposedly indicative of empathy. In the literature on the Rorschach test it is sometimes claimed that the human movement response (M) comes about only through an empathic, kinesthetic identification of the subject with what he perceives in the blot. Our experience has been that movement responses may be empathic in origin but are not invariably so. In an empathic M response the subject vividly recreates a postural experience or attributes thoughts and feelings to the imagined person in an appropriate way. In the TAT the treatment of the thoughts and feelings of the characters would seem to be a royal road to the measurement of empathy, but we did not find it to be so except at the extremes.

As demonstrated in Chap. 16, it is worth while to look for signs of Freedom from Ethnocentric Prejudice, both for its direct bearing on work with members of outgroups and for its value as an indicator of an undesirable authoritarian character structure.

It is not difficult to inquire into attitudes toward Jews, Negroes, or members of various foreign national groups in the interview, though the answers one gets cannot be taken at face value. Usually, if a person declares himself to be prejudiced, he is; but just how harmful the degree of prejudice may be is hard to say. More commonly one encounters lip service to democratic values without conviction. Frequently, the tone of protesting too much can be caught in such declarations, or a person may express stereotyped and subtly derogatory attitudes toward ethnic group members in the very process of claiming "tolerance."

It is relatively uncommon for evidence of ethnocentric prejudice to emerge in the usual psychological tests, so when it does, it is the more valuable. Thus, applicants occasionally see the male figure in card 15 of the TAT as a Negro, and it is only moderately unusual for the dark-haired girl in card 4 to be called a "mulatto," "Eurasian," or the like. The *attitude* toward members of such groups is usually

quite apparent in the story. Similarly, certain dark ink blots in the Rorschach test occasionally inspire applicants to identify figures as Negro, and derogatory attitudes are sometimes expressed in such a context by people who will stoutly maintain antiprejudicial attitudes when asked directly about it in an interview or in a test such as the California Ethnocentrism Scale. Certain popular percepts (*e.g.*, "cannibals" or "savages" for the popularly seen human figures on card III) do not indicate ethnocentric prejudice unless they are elaborated in a way that makes it unmistakable. In as sophisticated a group as applicants for psychiatric residency, such an instrument as the California E Scale is too crude and obvious to be of value (see Chap. 13).

The kind of generalized nurturant and humanitarian sentiments that we summarized as Consciousness of Social Injustice can best be assessed in much the same way as freedom from ethnocentric prejudice.

Interview discussions of social and political topics are obviously indicated, though the more that such material can be elicited subtly without direct questions, the more reliable is it likely to be. Reactionary, antidemocratic values on the one hand, and socially conscious ones on the other, are easier to assess with the TAT than is ethnocentric prejudice, since a number of the pictures, particularly 3 BM, 13 MF, 4, and 2, often give the impression that the persons pictured are economically deprived. Again, it is particularly valuable to compare direct and indirect indications of such value attitudes. Discrepancies sometimes indicate a lack of genuineness (see below) or self-objectivity, though sometimes they merely reveal conflict. The latter may be a healthy sign, as in the case of a person reared in the South with some emotional prejudices which he is honestly trying to overcome on the basis of a consistent philosophy of life.

**SOCIAL ADJUSTMENT.**—It is not necessary that an applicant have attained a Mature Heterosexual Adjustment for him to make an adequate, even superior, resident. But we found that attitudes toward sexuality, strivings toward sexual adjustment, and orientation toward the opposite sex were related to Over-all Competence and Psychotherapeutic Competence. If the applicant is oriented toward love, making efforts to find a mature relationship of mutuality and intimacy with a sex partner, this is an indication of a generally favorable sort. Unsatisfactory applicants frequently had had histories of avoid-

ing contact with persons who might make suitable mates; of homosexual orientation more or less marked in degree; of marked, but usually unconscious, hostility toward members of the opposite sex; or of exploitative, parasitic, sado-masochistic, or otherwise destructive relationships. These patterns should be distinguished from one of delayed sexual maturity, with dating only late in adolescence—or even not until college years—and a slow development of bodily sexual characteristics. This kind of late flowering of sexuality was not uncommon among superior residents and was frequently encountered in the total group of those whose work turned out to be adequate.

The subject of feminine identification and latent homosexuality is a complicated and controversial one in this context. Our impression is that psychiatric residents as a whole have a stronger degree of identification with their mothers and, accordingly, a stronger feminine component in their personality make-up than is true of men in general. If a man feels comfortable with, and accepts, his more feminine aspects, though still developing a basically masculine identity, this is all to the good. Similar amounts of feminine identification may also lead, however, to overt homosexuality or latent homosexual trends strong enough to disrupt relationships with both men and women; or to paranoid trends (see *Freedom from Projection*, below); or to a hypermasculine protest which, in turn, may disturb relationships and interfere with the functioning of such "feminine" qualities as sensitivity, empathy, and need to help others.

A few test signs of disturbance in sexual adjustment were particularly useful: destructive relationships between men and women in TAT stories; in the Rorschach, sexual organs seen in unusual areas that objectively bear little resemblance to them, frankly hermaphroditic human figures or figures seen as wearing clothing of the opposite sex, seeing the same area (or bilaterally symmetrical areas) as male and female, or seeing genitals in areas where they are most commonly seen but attributing the wrong sex to them. These types of Rorschach responses usually indicate severe conflict over sexual identity or pathological sexual preoccupations, especially when accompanied by indications that the subject is threatened by them, and tend to be found in the records of unsatisfactory applicants.

Assessments of this variable through the interview, on the other hand, usually failed to yield information of value in predicting psychiatric performance. Moreover, it should be borne in mind that clumsy, unempathetic, direct questioning about sexual experience and

practices can be unnecessarily traumatic to the applicant. Among the few complaints that came to us about the psychiatric interviewing for the Menninger School of Psychiatry, this was the most conspicuous. The interviewer should at all times remember that the situation is very different from the usual diagnostic or therapeutic one with a patient and that the candidate knows his confidences will, in a certain sense, not be respected.

Our variable, Social Adjustment with Co-workers, owes its clumsy title, and some of its usefulness as well, to its "double-barreled" character. On the one hand it points toward *specific* relations with professional and semiprofessional colleagues; on the other it directs us to look into the applicant's *general* level of social adjustment.

Drawing out the applicant on how he has gotten along with nurses and fellow interns gives useful information concerning the specific aspect of this variable. Questions on the kinds of difficulties he has had in dealings with nurses, followed up by requests for specific anecdotes, can be particularly revealing. In the TAT the relationships between heroes and their peers—siblings, co-workers, and friends—should be evaluated for the more general aspect. Stories in which the central figure characteristically has friends, deals easily with people who in turn are generally helpful and co-operative (without simply being depicted as subservient to the hero's desires or as wish-fulfilling cat's-paws) are strong positive signs. If a man has envious, overcompetitive, domineering, overingratiating, or coldly distant relationships to the people with whom he works, corresponding attitudes will often be mirrored in the stories. At times, these trends may be more clearly seen after the subject's attempt to interpret his own stories. Letters of recommendation from internships sometimes contain relevant information on this score.

By Feelings of Security we mean a benign view of the world as a supportive rather than a barren or threatening place.

The feeling of basic trust and optimism—that things will turn out all right, that people are generally good, that supplies will be forthcoming—may at times be seen very clearly in the TAT stories. One should be on the lookout, however, not to be taken in by over-pretty projected situations based on denial rather than security; sometimes the discrimination is not easy to make. Good residents frequently told stories of conflict and suffering, representing inner struggles through which they were going, that could easily be mis-

interpreted to mean a hopeless outlook or view of the world as a place full of frustration and suffering. It is most helpful to look for a mixture or variety of these kinds of content; as long as it is not wholly one-sided, the balance can swing fairly far toward either the grim or the gay. Self-interpretations sometimes help to clear up these points, but interview data, particularly on personal history and development, usually serve to give the proper balance. In the interview, too, one must distinguish carefully between reports of a childhood that was actually full of happiness and loving support from parental figures and bland cover-up stories stressing "normal" development and the absence of trauma in the effort to conceal what the applicant fears may be damaging admissions.

In retrospect, we feel that it might have been more useful to have conceptualized the closely related areas of self-confidence and feelings of security in terms of Erikson's concept of the *sense of identity*. (Erikson, 1953, 1956.)

**HANDLING OF EMOTIONS.**—When we turn toward Adequate Emotional Control, little needs to be said to the psychiatrist or clinical psychologist; he is on home territory. Indeed, all three of the variables included under the general heading of "Handling of Emotions" are generally evaluated by the psychiatric interview, used as a psychological examination. On the general topic of emotional control, however, see the above discussion of emotional warmth under Need to Help. The optimal balance between free expression of impulse and feeling, and their controlled, modulated channeling, seems to lie somewhat on the controlled side of the midpoint. Our observations were that the inexpressive, inhibited candidates were better able to loosen up adequately than the freely affective, impulsive ones were to learn adequate delay of impulse and control of emotional expression. As long as there is genuine interest and concern behind it, patients are less likely to be repelled by an inexpressive, over-controlled exterior than they are to be frightened or upset by too much display of affect.

Adequate Handling of Hostility and Emotional Appropriateness are specific aspects of emotional control which we found it useful to attend to particularly.

Psychological tests afford numerous direct and indirect opportunities for assessing all three of these emotional variables. Careful observations of the applicant's behavior in the testing situation itself,

as varying demands are made on him, as he is presented in the TAT with pictures having strongly sexual or aggressive connotations, or others that tend to arouse depressive or guilty feelings, and as he is shown the various heavily shaded or brightly colored cards of the Rorschach test—all of these situations afford opportunities to catch the candidate off his guard and see the quality of his emotional control and the appropriateness of his expression. In addition, the Rorschach test can often tell a good deal about emotional control through the way the subject responds to color and integrates it into his responses, though on this point we should like to warn that a great variety of patterns of response to color were found in the Rorschachs of adequate residents. A total absence of color responses, even weak or possible ones, was very rare among them, however, although not uncommon among inadequate residents. Other Rorschach signs characterizing inadequates were: a complete absence of FC (form-color) responses; the presence of C' responses (responses to achromatic color in which form played no part, such as "mud" or "snow"), especially if there were two or more pure C'; failure to give any response at all to brightly colored cards or marked delay followed by poor responses; and color-naming rather than real responses. These are all rather extreme and unusual reactions, but any of the more familiar types of response were found just as often among acceptable as among unacceptable applicants. In general, our experience in the assessment of emotional control variables with psychological tests agrees with that of Rapaport *et al.* (1945-46) and Schafer (1948, 1955).

One further word about Emotional Appropriateness: although this variable is most important at its lower extreme, when emotional reactions grossly inappropriate to the situation indicate schizoid disorganization (see below) or psychopathic blandness, we found it useful to distinguish milder degrees of failure to perceive the affective nature of the situation quite accurately, or to react in a way that was wholly in keeping with the implicit demands of the situation. These less clinically obvious degrees of inappropriateness seem to indicate a failure to attain the level of differentiated emotional experience and awareness that is necessary for following the subtleties of emotional communication from patients, and reacting to them sensitively and with tact. The TAT is useful in this kind of assessment; emotional inappropriateness can be diagnosed by a failure of the stories to develop with smooth emotional logic, as with

sudden and forced happy or unhappy endings, or other arbitrary turns in the story content.

**INTELLECTUAL EFFECTIVENESS.**—The first successes of psychological testing were in this area, and it remains one in which tests must be relied on heavily. Among the qualitative features of intellectual performance that can be gauged are Clarity of Thought and Freedom from Stereotypy of Thought. In our Predictive Study these turned out to be among the less useful variables to look for, but they still have some value at the lower extreme: unclear and stereotyped thinking are more characteristic of unacceptable than of acceptable candidates. We were surprised, however, to find so much of both in our general population. Crystalline lucidity of thinking and creative imagination are desirable but apparently not necessary qualities in a psychiatric practitioner; they would take on more importance if one were laying particular stress on training teachers and researchers in psychiatry.

The interview can be used to estimate clarity of thought, although it takes a sharp and trained ear to detect vague or muddled thinking if it is glossed over by a pleasant manner and a free flow of words. We found it useful to draw applicants out on psychiatric concepts: Those who were low on clarity of thought would often show a kind of pseudo-thinking, a mishmash of technical terminology without any attempt to meet issues squarely. Ignorance or misinformation should not be held against a man who has not yet studied psychiatry, if simply admitted or not compounded by the attempt to discuss complex issues without really understanding them. In the Self-Interpretation, clarity of thinking shows itself in an attempt to be systematic, well-organized, and sensible rather than chaotic, vague, or empty. Stereotypy of thought may be judged by the proportion of animals in the Rorschach (65 per cent or more generally indicates banal, cliché-ridden thinking) and by the originality of TAT stories: are they fresh, with at least some individual twist, or are they the obvious plots embellished only with cant phraseology and sentiments?

The most useful recommendation we can make about Perceptual Sensitivity is not to bother with it. As gauged by tests, it characterized a few excellent residents, but fair-to-poor perceptual sensitivity was so well distributed throughout the range of talent that it did not distinguish acceptable from unacceptable applicants.

The kinds of sensitivity that are most relevant to psychiatric diagnosis and treatment (often spoken of as observational ability, social sensitivity, or even intuition) are probably taken care of adequately by assessments of psychological-mindedness and empathy.

Another in the list of nonessential but desirable qualities in a psychiatric resident is Cultural Wealth. The low correlations between ratings of cultural wealth and our criteria suggest that it may be overvalued in some quarters. It is possible for a man to become a good psychiatric resident even though he has come from a background in which there was nothing to foster an interest in art, music, poetry, philosophy, or the other humanities. But if a man has had as much opportunity as most of his fellows to be exposed to these aspects of culture and has not developed taste, familiarity, and above all a genuine interest, there is some reason to question his suitability as a psychoanalytic psychotherapist. (For discussion of this variable see Chaps. 15 and 16.)

It is worth while to pay some attention to Quality of Verbalization, but it should not be given heavy weight.\* Indeed, this variable is of little significance unless the candidate's verbal facility is markedly low. As was true of Cultural Wealth, we occasionally encountered men who were very gifted in their use of words and who made very unsatisfactory residents. The important considerations are, first, how well a resident will be able to communicate with patients, and second, how good a job he will do in communicating with colleagues, particularly in conference reports and written diagnostic summaries. One occasionally encounters a relatively inarticulate man who nevertheless succeeds in getting across his points unmistakably, especially in his relations to patients; such people should be carefully distinguished from those who use language insensitively, awkwardly, and stumblingly without compensating gifts for nonverbal communications.

Quality of verbalization can be judged in any interview or test situation; fine discriminations about it are easiest when one has the verbatim record of what the subject has said.

**HONESTY.**—In approaching the problem of Well vs. Poorly Internalized Ethical Standards, we focused on its latter aspect, which is easier to assess.

\* The Verbal IQ, which correlated well with competence (Chap. 13), measures more conceptual types of verbal ability, not skill in using words.

On looking back over the initial assessment of men who later showed deficiencies in integrity, we find that interviewers frequently developed a feeling of distrust for such men but too often were disinclined to listen to their intuitions. The interview seems to be the most sensitive instrument for picking up the subtle indications of poorly internalized ethical standards. Oftentimes, however, the most important evidence comes from inconsistencies in the story told by the candidate to different interviewers, or discrepancies between his verbal account of events in his past and the documentary record that may appear in credentials. Certainly there should be careful interview inquiry whenever a letter from a former teacher or work supervisor raises any question about integrity.

Perhaps the most frequent manifestation of a lack of integrity is opportunism; it may be detected in the work history, reasons for wanting to enter the psychiatric profession, attempts by the applicant to shift his sails with what he perceives to be changes in indications of the interviewer's interests or values, and the like. Sometimes lying and evasiveness can be picked up from internal inconsistencies, but it is particularly useful for the various interviewers to compare the details of the candidate's story if there is any question about his integrity.

In psychological tests, an inclination to cut corners on the test requirements can at times be picked up in the Wechsler-Bellevue. Certain unnatural patterns of Rorschach responses, in which types of response often considered "good," such as human movement responses,\* are unnaturally frequent, should occasion careful inquiry about the candidate's previous experience with the test or reading about it. Anyone who is not very thoroughly familiar with psychological tests will find it extremely difficult to falsify his test picture in a supposedly ideal direction, though it is sometimes tried. Finally, dishonest and opportunistic behavior on the part of story heroes sometimes appears in the TAT's of applicants with psychopathic trends. Avoid jumping to unfavorable conclusions if such impulses have been successfully overcome in stories, or if they appear in the context of an attempt to be "clinical" or understanding about all varieties of human behavior.

One of the most predictive individual variables of personality

\* Dr. Roy Schafer, in a personal communication, suggests that "artificial color responses like 'blue hat,' and 'red chair,' etc." often raise the question of the candidate's prior familiarity with the test.

was Genuineness vs. Façade, yet considerable care is necessary in assessing this variable. The main problem is to distinguish between characterological distortions and situational defensiveness. It was rare that any of our predictive judges were taken in by a façade, but it not infrequently happened that we overevaluated the significance of pretentiousness and pose in men who later turned out to be very good residents. In the latter cases the stress of being evaluated caused basically honest and genuine people to try too hard to put their best foot forward. This distinction is not at all easy to make, except on an intuitive basis from direct contact with a man in interview.

Phoniness can show up in a myriad of ways, either in interviews or in tests. A pompous or blustering manner, covering inner insecurity, is obvious in some applicants whenever they open their mouths; others become pretentious mainly when they are called on (as in the Wechsler-Bellevue) to display their intellectual powers and attainments. Most serious is a falseness of feeling that goes deep enough to be apparent in the projected reactions of story heroes in the TAT. Another way in which pervasive characterological distortion can be distinguished from a situational reaction is through the consistency with which indications of façade turn up in various interviews and tests.

PATHOLOGICAL TENDENCIES.—Evidences of overt psychosis, addiction, perversion, psychopathic character disorder, or other serious psychopathology is usually enough to warrant rejection. In the Admissions Committee of the Menninger School of Psychiatry careful inquiry was always made as to whether the illness was current, whether sufficient treatment had been received, whether past performance had been markedly impaired, and whether the applicant had special counter-balancing talents that made it worth while to take more of a risk with him. Little needs to be said about techniques of assessing psychopathology because the psychiatric interviewer or the diagnostic tester is trained specifically to look for evidences of it. We may comment, however, that one should not strain at gnats in probing for pathology. It seems to be particularly true in diagnostic testing that many clinicians are apt to make too much of minimal test indicators, which may be massed in the records of men who turn out quite adequately as psychiatric residents and may even be potential assets as a basis for understanding and empathy. It is by now common knowledge that well-functioning, apparently healthy, people may have a

good deal of encapsulated pathology that can be kept separate from, and in some ways even facilitate, their everyday functioning. One should not lose sight of common sense considerations, therefore, when weighing the significance of test signs of paranoid trends or occasional schizophrenic-sounding responses. It is good to have the work history (from credentials and interview) in mind when evaluating psychopathological indications from interview or tests; and if they do not look really malignant, one should not be overly concerned about pathological tendencies in a man who consistently has been operating effectively.

A special word of caution may be in order regarding the Rorschach test for anyone who has not had considerable experience in giving it to intelligent persons who are not seeking treatment. An applicant often perceives the Rorschach situation as one in which he ought to let himself go as much as possible, give free-associative material, or demonstrate the liveliness of his imagination by what are essentially primary process operations in his perceptual-associative functioning. The result often bears a superficial resemblance to schizophrenic productions: There may be extensive autistic elaborations, arbitrary combinations or fusions of images, or a massing of sexual content. If the tester has any doubt, he can quickly establish the true nature of these productions by readministering the test with instructions to leave out that kind of material; when it is truly pathological in origin, the subject cannot long exclude it from consciousness. Internal checks on the significance of such material, whether it represents pathological break-throughs or regression in the service of the ego, can be seen in the preservation of formal properties, such as the accuracy of form perception, the balance of locations and determinants, and the like. (For further discussion of these matters see Schafer, 1955; and Holt, 1956.)

#### SPECIFIC RECOMMENDATIONS TO SELECTORS

A dozen conclusions follow which summarize the major implications of this research project for the practical task of choosing among applicants for psychiatric residencies.

1. *The final decision on accepting or rejecting a candidate for residency training is best entrusted to a multidisciplinary committee of several members, operating under democratic leadership and con-*

*centrating their discussions on moot or difficult problems.* During the years that the Menninger School of Psychiatry has been in operation, a number of different working procedures have been tried out in the successive Admissions Committees. Observations both from the sidelines and from our participation in some of these committees are the source for the following remarks on ways that such a committee might operate.

The members of an Admissions Committee may themselves not be directly involved in interviewing or testing applicants; they may deal entirely with reports and recommendations handed to them by other persons. Whatever may be gained in objectivity through such a procedure is more than compensated for by the loss in information; it seems best to have as many members of the committee as possible directly involved in testing and interviewing.

A committee may be run in a variety of ways, depending on the skill, outlook, and preferences of its chairman; different purposes and different settings make a more democratic procedure efficient in one case, a more autocratic one preferable in another. Work of the present kind proceeds most expeditiously in a democratic group atmosphere; no one person can be assumed to have special competence in judging all kinds of applicants.

There are various ways in which the three principal sources of information (interview reports, test reports, and credentials) can be used by a committee. No matter how good the individual components may be, if they are not used properly their value may be wasted. A procedure often followed is to disregard the psychological test report or letters of recommendation when they are markedly at variance with the interviewers' consensus. We have reported in Chap. 9 that in cases where psychological tests and interviews disagreed, the judgment of the tester was in fact usually better, but we would not advocate following so mechanical a method of resolving such differences. When there is a discrepancy between two different sources of information, this points to a problem that needs to be resolved by a synthesis; simply to decide that one source of information on an applicant is correct and another incorrect is to overlook the possibility of a real insight broader than could be achieved by either of the disagreeing approaches. Indeed, applicants about whom one's sources of information disagree are the ones that require the judgment of a committee. Little is served by taking up

the time of committee members in hashing over cases on which there is universal agreement that the candidate is clearly acceptable or unacceptable.

The most efficient procedure we have observed operates in the following fashion: First, a file folder is made up for each eligible applicant who completes the process of assessment. Into it go his application forms, reports of medical examination, letters of recommendation, and similar material that we refer to under the common heading of Credentials. Reports from interviewers and psychological testers are similarly filed in this folder, and when the information is complete, it is circulated to the various members of the Admissions Committee. Each member studies the contents of the folder, records a definite vote or a recommendation to accept or reject, together with whatever qualitative notes or conditional remarks that he may wish to make. This procedure is carried out independently by each member of the committee; no member sees the judgments and comments of the others before recording his own. When the folder comes back to the chairman, he then collates the recommendations and reads the comments. If the committee is unanimous in recommending acceptance or rejection, and if there are no serious qualifications noted by any of the members, the chairman then takes definite action on the case, and no committee meeting is called. When there is disagreement, or when the comments of any member indicate that he has serious doubts or feels that something would be accomplished by group discussion of the case, it is set aside for discussion at a committee meeting, which is held whenever enough such cases have accumulated to warrant a meeting (or sooner if there is particular time pressure).\*

In such a meeting, the chairman presents the basic facts on which there was disagreement and either follows this by a synthesizing interpretation of his own or invites the members of the committee to contribute their ideas. The discussion is focused on putting the various pieces together into a consistent picture of a human being acting differently under different circumstances, or being viewed differently under the various lenses of the procedures used. With the guidance of the chairman, the group attempts to achieve a real consensus, decision by majority vote being resorted to only in case of

\* It is good to notify applicants as soon as possible after they have been seen. The best prospects usually have other opportunities, and they may decide to go elsewhere if the decision is slow to reach them.

impasse. The free interchange of ideas and hypotheses by a group of men as many of whom as possible have had direct contact with the candidate usually produces an intelligible personality picture, leading to a decision with a good chance of validity.

How should such a committee be composed? In the Menninger School of Psychiatry, the Admissions Committee has always included a psychologist and often an educational administrator, along with a preponderance of psychiatrists, all of whom are members of the panel called upon for admissions interviewing. Sometimes the psychologist member has done all of the testing; at other times he has done none of it. At any rate, it seems advisable to have at least one discipline other than psychiatry represented on the Admissions Committee.

2. *Best results are achieved by finding interviewers and testers who have a flair for selecting residents.* There is some evidence that the ability of the person making predictive judgments is important, more so than the particular techniques that he uses (see Chaps. 9 and 12). Consequently, it is worth while to undertake the long task of trying out a group of interviewers and testers, following up the success of their predictions upon as many applicants as possible and with the best criterion data obtainable. The temptation to use agreement between judges as a cue to competence should be resisted; suitable criteria are suggested in Chap. 8. Once a good predictor has been found, every effort should be made to withstand pressures of the kind described above to get him to cut down the amount of time he devotes to Admissions Committee work. It is a false economy to skimp the assessment of applicants and then waste months of valuable training efforts on unsuitable candidates.

3. *If you can afford only a minimal procedure, combine a careful evaluation of application credentials with psychiatric interviewing.* The method of evaluating credentials expounded in Appendix 10.7 together with a psychiatric interview seems to us to be the best compromise between the implications of our research and the practical considerations mentioned earlier in this chapter.

4. *Where a more expensive procedure is possible, use a well-rounded battery of psychological tests administered by a skilled psychologist and then increase the number of interviewers up to three.* If something more ambitious than a single psychiatric interviewer is possible as an assessment procedure, we would advocate first adding psychological testing and then further interviewing.

Though the average of two interviewers' judgments was better than one, and three better than two, the added validity contributed by a fourth interviewer would not be worth the additional cost. Considerations of economy should not be allowed to reduce psychological testing to the administration and interpretation of a single instrument, such as the Rorschach. If testing is done at all, it should be done right; our results suggest that even a group of projective tests need to be supplemented by nonprojective ones in a well-balanced battery.

5. *Provide for the continuous training of assessors through a systematic feedback of results of their predictions.* Routinely, every interviewer and psychological tester should be informed how well he has been doing as soon as criterion data become available. At times, they should be presented with names of all applicants about whom they have made predictions so that they can go back over their reports and notes and learn as much as possible from their successes and failures. Though we have no research results to back up this recommendation, there is reason to believe that this device should improve the accuracy of the judges.

6. *Do not give much weight to grades in medical school unless an independent study in your own setting reveals a definite relationship.* A careful study of grades in medical school failed to turn up any relationship with performance in psychiatric residency (Chap. 13). A survey of current selection practices shows that this information is given a great deal of weight in many institutions; therefore we call special attention to it.

7. *Other things being equal, an applicant who is beyond his middle thirties, or older, should be considered with special caution.* In Chap. 13, we have presented evidence supporting the view that the younger applicants make better psychiatric residents.

8. *Reports to the Admissions Committee by interviewers and psychological testers alike should contain a full qualitative discussion of the applicant's assets and liabilities, a definite recommendation to accept or reject wherever possible, and full discussion of reasons if a "doubtful" or undecided recommendation is given.* It may be useful also to have the assessors make a quantitative rating on a clearly defined rating scale. It seems best to leave the style and contents of the qualitative report largely up to the individual writing it, but a short outline of topics ordinarily covered may be helpful, particularly to interviewers and testers just beginning to do the

job. Time pressures will exert a steady temptation to minimize qualitative discussion, particularly when the decision seems very clear-cut, but this should be stoutly resisted. In case of disagreement within the committee, in case the interviewer or tester cannot be present at the discussion of the applicant, or in case the assessor is reviewing his actions some years later in an attempt to understand and improve his own procedures—under these and other similar circumstances, full qualitative notes are of foremost value.

9. *An applicant with a Verbal IQ below the Superior range (119 or lower on the Wechsler scales) should be admitted only after special consideration; an applicant with a Verbal IQ above 136 has a good chance of making a superior resident.* The Verbal IQ from the Wechsler-Bellevue Test of Adult Intelligence maintained a significant low positive correlation with performance as a resident through our study, and therefore ought to be considered in selection (see Chap. 13). We did not try the new revision of this test, the Wechsler Adult Intelligence Scale, and do not know how well it would do the same job, but it seems well worth trying, at least in selection research. In addition to the usefulness of the IQ itself, the Wechsler tests also afford rich opportunities to observe the style and quality of intellectual performance. From a verbatim record of the candidates' responses, a skilled psychologist can make useful estimations of the extent to which a man is using his native endowment, the nature of his working habits, his capacity to maintain the efficiency of secondary process functioning in spite of the presence of neurotic difficulties, impulsiveness vs. judiciousness of judgment, breadth of information, and other matters of relevance.

10. *A careful evaluation of an applicant's past performance in psychiatric work is one of the most useful predictors of success as a resident.* Our findings suggest that it would be useful to get as much information as possible about psychiatric clerkships and the psychiatric part of a rotating internship, as well as evaluations from the supervisors of those men who already have had some residency training.

11. *When the applicant cannot reasonably be expected to appear in person for evaluation, a suggested second-best procedure is to have him send information by mail.* The following battery of procedures has been used with some success at the Menninger School of Psychiatry, especially for applicants from distant countries whose facility in English it is important to assess:

### 332 Conclusions and Recommendations

- a) Autobiography (with instructions about length and areas to be covered).
- b) Vocabulary test.
- c) Attitude questionnaire (a group of open-ended questions dealing with attitudes toward parents, psychiatry, and other topics).
- d) Signed statement from a competent authority attesting to degree of facility in English.
- e) In addition, the usual credentials described in Chap. 10.

12. *If you are choosing ten residents a year or more, set up a continuing research project into your own operations.* In the long run, everyone's interests will be best served if a modest program of operations research is undertaken by each psychiatric training institution. A relatively simple procedure, such as the one described in a suggested research of this kind for analytic institutes, could provide a great deal of useful information (see Chap. 18).

### RESEARCH RECOMMENDATIONS

In over eight years' work on this project we feel that we have learned a good deal about ways in which large-scale research upon the selection of professional personnel can best be done. We are presenting here a set of recommendations, for the guidance of anyone who wishes to undertake a project similar to ours.

1. *Allow plenty of time for careful planning.* For a project that is to last five or more years, a full year or more of detailed planning and pilot studies can in the long run become the best way of preventing wasted time. Free use should be made of consultants in all phases of the subject matter to be studied, as well as in all aspects of research design and the treatment of data.

2. *Make sure that lines of responsibility and authority are clearly laid out and understood by all members of the research team.* Research in the selection or education of psychiatrists, psychoanalysts, or similar professional people necessarily takes a long time and a collaborative group of investigators. Co-operative research is difficult under the best circumstances, but particularly so if attempted without a clearly articulated group structure. It is equally important to give continuous attention to personal and professional relations between the research group and cooperating members of the profession under study. The difficulties of communication between re-

searchers and practitioners are both many and subtle; smooth working relationships require constant effort.

3. *At least one person in the research group, and preferably the one with major responsibility, should be able to devote full time to the project.* It is tempting for people whose major responsibilities lie in teaching, clinical practice, administration, or some combination of these, but who have a moderate amount of time for research, to think that they can run a large-scale selection research. No doubt this has been done occasionally with success, but it is safer for those with the heaviest responsibility to have most of their time committed to the project. Moreover, anyone who is expected to carry out a major share of the work on a project must have a real part in its planning.

4. *It is of paramount importance to consider the objectives of a study with the greatest care.* The scope of the work to be done needs to be restricted in the most hardheaded way, in terms of the actual resources of money and time that are available. In planning research, particularly in an applied field, it is always worth while to perform the mental experiment of supposing that the answer to the first questions posed have been obtained and to ask oneself, "So what?" It very often happens that when a problem is not clearly defined and based upon real needs, some aspects of it evaporate quickly under this treatment. Resist the seduction to add subprojects simply because it would be "interesting" to have the results, if they are not clearly relevant to the major underlying objectives.

5. *Selection research should begin with a thorough job analysis.* In order to predict whether a man is going to be a good practitioner of some profession, one needs to know in considerable detail just what the practice of that profession entails. This means, ideally, at least a number of months devoted to empirical studies of what the professional does, and the isolation of concrete instances of behavior considered good and bad practice (on the method of critical incidents see Flanagan, 1954). This leads directly into a closely related problem, that of the criterion.

6. *Intensive work on the criterion needs to be done at the outset of any selection project.* Even if a job analysis has been adequately carried out, it still does not make sense to begin predicting until one knows how the behavior in question is going to be evaluated—in short, just what it is that one is trying to predict. In select-

ing trainees for one of the learned professions this is a particularly difficult problem, since such persons often work in relative isolation and conventional criteria of productivity are generally not available. It may be difficult to determine just how early in a man's career it becomes possible to see the quality of performance that will characterize his career as a whole. In this connection it may be desirable to make a special study of the careers of a few prominent practitioners about whom there is very widespread agreement on their competence, with the aim of finding out at what time this competence became manifest and in what ways.

*7. Make certain that a large enough group of subjects will be available within the period of time the research group is willing to commit itself to work on the project, in order to yield significant results.* Our experience strongly suggests that a single cross-validation is not enough; there may be enough sample variation from one year to the next in the kind of persons seeking to enter a profession to make it necessary to try out any predictors on several successive groups of applicants.

*8. Explore new techniques and instruments for selection.* Letters of application and recommendation, application forms and the like—constituting what we call Credentials—have been generally considered of negligible value in selecting members of such professions as psychiatry, though they are almost universally used. Our experience suggests that a systematic study of such materials from applicants who later on turn out to be effective, or ineffective, may lead to an assessment technique of considerable value. This is a type of research that needs to be pursued more intensively than it has been in the past.

The manual method with projective techniques such as the Rorschach and TAT showed enough promise to suggest that further work, building on our findings, might ultimately yield useful instruments. Cues that retained a modicum of validity and reliability throughout our Small Sample Study and both cross-validations could be supplemented by the development of others and the continual refinement of manuals to promote the reliability of scoring. In any such future developmental work with TAT, the cue we called "adequacy of the hero" appeared promising and is worth more research effort.

Despite the fact that the few inventories and similar instruments

we used turned out poorly, recent developments in this kind of testing suggest that it should be given a further try. Some scales that have been developed for Gough's California Personality Inventory and the Minnesota Multiphasic Personality Inventory (such as Barron's "ego-strength" scale; see Barron, 1953) appear promising.

An information test covering knowledge of psychiatry could easily be developed and, when administered to applicants, might provide a useful measure of interest in the field. This is based on the assumption that differences in the amount of knowledge of a field *before* formal and intensive instruction probably reflect differences in interest that have led one man to read more on his own and remember more psychiatry from undergraduate medical education than another who cares less. This assumption seems to be validated in studies on the recognition of aircraft (Luborsky, 1945) and the selection of aircrew personnel (Rohles, 1953).

Situational tests, the use of standard role-players, group interviews, work-sample tests, such as having applicants interview patients while being observed—these and many other techniques deserve trial, which we were not able to give them.

9. *Experimental selection methods should always be presented to the applicant as if they were going to play a significant role in his acceptance.* Even though this involves a mild amount of deception, it is necessary so that the subject's attitude toward the experimental test should be the same as it would be in possible operational use later on. The same point applies to preliminary studies of extreme groups, such as our Small Sample Study: such undertakings are most useful if they are confined to data that were gathered at the time of application rather than later on when the degree of competence has been established.

10. *Intensively study selected individuals throughout their training,* choosing good and poor prospects, and prospects who are good or poor for a variety of reasons. Note changes in their evaluations of self, supervisors, colleagues, patients, and the field they work in. In this way, keep in touch with (a) intimate details of student life and experience, (b) beneficial and harmful experiences or relationships, (c) the effect of changing life circumstances, etc. A near-therapeutic type of relationship between researcher and subject may develop, but skill and reserve may regulate this effectively. The steady review of experience with an interested and appreciative per-

### *336 Conclusions and Recommendations*

son may itself have a salutary effect, but this can be controlled by comparison with the development (in ratings) of an equated group of residents.

11. *Try to have as stable a criterion situation as possible, i.e.,* well-established training procedures, relatively little turnover in supervisors, interviewers, and testers, etc. Obviously room must be left for growth of the training organization; ideally, however, it should not be neonatal growth but that of early maturity.

12. *Do not neglect to follow up persons who are rejected for training in the institution or institutions where the study is being done.* Even the data that can be obtained from mailed questionnaires may be very useful in establishing criterion groups and in not wasting a substantial proportion of the subjects who are assessed.

APPLICATIONS TO THE SELECTION  
OF CANDIDATES FOR  
PSYCHOANALYTIC TRAINING

IN THE CLOSELY RELATED FIELD OF PSYCHOANALYSIS, THE PROBLEM OF selecting new members of the profession has been becoming steadily more salient over the past decades. Psychoanalysis began as a small band of pioneering, even rebellious, spirits who gathered around Freud; as their numbers grew and centers of practice spread outside Austria and then Europe, the position of the group among the disciplines related to medicine changed slowly. For a long time, opposition to psychoanalysis was so vehement and so emotional that only a dedicated few sought out the opportunity to become analysts. Selective procedures were informal, even after the decision at the end of World War I to adopt Nunberg's proposal that new members should first be psychoanalyzed themselves. The first institute for the training of analysts was founded in Berlin in 1924; during the decade of the twenties, others were set up, and the number of applicants to these institutions gradually increased.

As time went on, it became necessary to formalize selection procedures. In all institutes the principal method has been interviewing, though candidates have usually been required to submit documentary credentials of the kinds we have been dealing with in this book. Selection methods in psychoanalysis were much the same as in psychiatry, except that interviewing was more protracted and assess-

ments more searching. The greatest single difference was the preparatory analysis itself, which was not only a therapeutic and didactic method but also a final screening.

The end of World War II saw a sharp rise in the numbers of candidates for psychoanalytic training, swamping the facilities not only for training but for selecting promising recruits. Up to that time there had been so few applicants in relation to the need that the task had been mainly one of screening out those who were clearly unfit (on the basis of either inadequate training or gross pathology). Now, however, for the first time there were more people who would have passed the old entrance requirements, more than the available time of training analysts could possibly accommodate. A Conference on Post-war Problems of Psychoanalytic Training was called, therefore, to discuss the situation. Representatives of all the institutes associated with the American Psychoanalytic Association met at the New York Psychoanalytic Institute on February 16 and 17, 1946, and reviewed all aspects of the problem.

In the deliberations of that meeting there were comments of a kind that have been echoed many times in subsequent years: not only had the quantity of applicants changed but also the quality. The arrival of psychoanalysis as a respectable and attractive specialty, with analysts being viewed by increasing numbers of psychiatrists as their professional elite, caused its appeal to become more widespread. Some participants in the conference believed that as organized psychoanalysis became an increasingly conservative body and its theory more popularly accepted, it would cease to attract men of an original and adventurous turn of mind. When a militant heterodoxy becomes a vested orthodoxy, it is bound to interest men of a different stamp. Some welcomed this change; others feared that it might cause psychoanalysis to lose its investigative birthright. All agreed, however, that it brought new problems in selection. The conference ended with the affirmation of a firm intent to maintain standards of selection and training, but without adopting any novel proposals for attaining these ends.

Another reflection of concern with the problem was the beginning of systematic research in psychoanalytic selection. To the date of this writing, no quantitative studies have been completed. The research being carried out at the Columbia Psychoanalytic Institute has made the most progress, although when it was reported (H. Klein, 1953) only one year of follow-up on a relatively small group of candi-

dates was available. Even less by way of validity data was included in a report to the Topeka Psychoanalytic Society by Ehrenreich (1955) on the rationale for the use of psychological tests in selecting candidates. (Contrary to the Columbia finding, Ehrenreich reported that the independent conclusions of the psychological tester were in "remarkably high agreement" with the decisions of the Education Committee to accept or reject an applicant.)

#### METHOD OF SELECTION

Detailed procedures for selecting psychoanalytic trainees differ from institute to institute, but the methods of the Topeka Institute are probably representative (except for the inclusion of psychological testing). The applicant fills out an application blank, which, besides calling for a usual curriculum vitae, asks questions on the origin of interest in psychoanalysis. At the same time, he writes a five-hundred-word autobiography. These materials are transmitted to two training analysts, who interview him after reading them. Each interviewer sees him at least twice, sometimes for as much as four hours in all; then forms are filled out and supplemented by free summary statements about the applicant. The great majority of applicants are, or have been, residents in the Menninger School of Psychiatry, which transmits to the Institute a file on each applicant at his request. It contains the reports of the interviewers and the psychological tester who saw him at the time of his acceptance into the School, together with reports on his performance in the residency. The test report is read by a psychologist who specializes in testing for the Institute, after which he administers another battery of psychological tests. (Occasionally he may request the file of the original test protocols to make a study of changes in certain test performances.) The psychologist writes a description of the personality organization and functioning in a good deal more detail than was done in the brief reports to the School of Psychiatry and includes a clinical judgment of the applicant's "analyzability": his capacity to profit from a psychoanalytic experience.

All of these data are studied by the training committee. Finally, the applicant has a group interview, with either the committee as a whole or a subcommittee. Further discussion takes place afterward, and a written decision is made along with a brief summary for the Institute's files, which may deal with priorities, special elements of

risk involved, etc. Some applicants are encouraged to change their plans and withdraw application in the course of the interviews. And, as one member of the Institute has written, "Selection simply means acceptance for trial training, so that it is true that the personal analysis, the didactic training as well as his controls, are a part of the whole selection process—lasting from four to five years" (Ekstein, 1956).

Problems of selection and training have continued to be matters of interest, concern, and even dissatisfaction among analysts. A full-day panel discussion of methods of selecting candidates for training was held at the midwinter meetings of the American Psychoanalytic Association in December 1953. Representatives of the different institutes described their selection techniques. The panel then considered the results of the Columbia study (H. Klein, 1953) and the Menninger Project (Holt and Luborsky, 1954; and Luborsky and Holt, 1957).

In the discussions at this and other professional meetings, many representatives of institutes have expressed the feeling that their procedures left much room for improvement. Some analysts have emphasized the lack of time to do the thorough initial assessment they considered essential; others have declared that a few observers are not enough of a basis on which to make so subtle a judgment. The whole training period is considered by most analysts to be also a process of mutual evaluation in which the candidate assesses his own capacities and interests, and the institute does the same (for a similar view *cf.* Ekstein, 1950). Some training analysts have deplored the lack of maturity and mediocre level of talent which they perceive in many applicants. In condemning the tendency in some institutes to submit applicants to ordeals of self-revelation through individual and group interviews, autobiographies, tests, etc., one training analyst was reminded of Freud's story about the medieval king who had an infallible way of telling whether or not a person was a witch: boil the suspect in a cauldron for three hours, whereupon the taste of the broth is a sure sign.

#### GENERALIZING FROM RESEARCH ON PSYCHIATRISTS TO PSYCHOANALYSTS

It may be harder to predict competence in psychoanalysis than in psychiatry because the members of the former profession are even

more highly preselected than are psychiatrists, and the therapeutic effect of the candidate's personal analysis introduces a large unknown. On the other hand, the criterion of competence as a psychoanalyst might be easier to predict. Since it is a matter almost exclusively of proficiency in psychoanalytic therapy, the criterion is more unitary and less diffuse than the one for psychiatry. If, on the other hand, the horizons of psychoanalytic selection were wide enough to put a special premium on persons who could become good teachers, researchers, or contributors to psychoanalytic theory, even this advantage might disappear.

On the whole, the problems of selecting psychoanalysts are strongly reminiscent of those in psychiatric selection. If the findings of the present project can be generalized to apply to any other professional group, it is to the psychoanalysts. The majority of the latter are drawn from the first ranks of psychiatric residents, and the psychoanalyst's work has much in common with that of the psychoanalytically-oriented psychiatrist. We have also seen in Chap. 15 that there is a very great area of overlap between the personality requisites for training in psychiatry and in psychoanalysis, as judged by experts in the two fields.

We have evidence that bears on two of the points made above: (1) that there is similarity in the requisites for work in the two fields; (2) that the applicants (or *accepted* applicants) are a select group from among the psychiatrists.

The evidence derives from the fact that most applicants to the School of Psychiatry were also judged during the screening interviews on their suitability for psychoanalytic training. The three-category ratings of "Take," "Doubtful," or "Reject" for psychoanalytic training were almost uniformly made during the School of Psychiatry's first two years, when the initial interviewers included the training analysts of the Topeka Institute for Psychoanalysts.

The interviewers' judgments of promise for psychiatric residency and for psychoanalytic training showed a great deal of similarity. For example, when Dr. A said "Take" or "Take?" for psychoanalytic training (in thirty-nine cases), he always said "Take" for psychiatric residency. However, he recommended accepting for psychiatry 17 per cent of those he rejected for psychoanalytic training. If the initial interviewers recommended "Reject" for a psychiatric-training candidate, they invariably recommended "Reject" for psychoanalytic training.

These figures suggest not only that there is a considerable overlap in the requirements for psychiatric and psychoanalytic selection but also that the standards for psychoanalytic training are definitely higher. Of course, this finding may be partly a function of scarcity—fewer openings for training are available in psychoanalysis.

Since not every psychiatrist who wants to be trained in psychoanalysis can be accommodated, institutes would be expected to choose candidates from among the most competent psychiatric residents. There is a trend of this kind among candidates in the Topeka Institute for Psychoanalysis, though less of one than we had expected. As of the closing date of our follow-up (July, 1953), 92 men from the research group of 242 residents had been psychoanalytic candidates. Within this group of 92 candidates, 29 per cent had been judged superior as residents—only slightly above the chance expectancy of 20 per cent. Similarly, those rejected \* for psychoanalytic training had been only slightly more often considered borderline or inadequate residents than had the residents as a whole. Also, a somewhat higher percentage of those who were rejected by psychoanalytic institutes were rejected by, or withdrew application from, the Menninger School of Psychiatry than was true among those who became candidates.

The trends look stronger if we break down the 92 candidates into "accepted by the Topeka Institute" and "accepted elsewhere" (Table I-18.1). Obviously the Topeka Institute was more attentive to the record of its applicants in the local School of Psychiatry than were other institutes. Of course, some of the most outstanding residents in the Menninger School of Psychiatry applied to the Topeka Institute as their first choice and, being accepted, naturally did not apply elsewhere.

Since psychoanalytic education takes many years, very few candidates have completed training. As of 1957, we had heard of the progress of 34 candidates in terms of definite completion, nearing completion, probable incompleteness, and definitely giving up completion of training. Assuming that these 34 are a fair sample of our group, we correlated their status with their earlier performance in psychiatric residency. The coefficient is low positive but not quite significant, indicating that the relationship, although in the expected direction, is less than one might expect.

As a postscript, it is interesting to compare the recommendations

\* This figure may not be exact—it was more difficult for us to establish that a man was rejected than that he was accepted by an institute.

TABLE I-18.1. PERFORMANCE IN THE MENNINGER SCHOOL OF PSYCHIATRY  
OF CANDIDATES AT THE TOPEKA INSTITUTE OF PSYCHOANALYSIS  
AND OTHER INSTITUTES

PERFORMANCE AS A PSYCHIATRIC RESIDENT (With Proportion of the Sample in Each Category)	TOPEKA INSTITUTE (During Residency)	TOPEKA INSTITUTE (After Residency)	OTHER INSTITUTES (After Residency)
Superior (21%)	65%	42%	10%
Adequate (60%)	35	58	70
Borderline (7%)	0	0	6
Inadequate (10%)	0	0	14
Drop-out (2%)	0	0	0
Total	100%	100%	100%
Number of cases	17	26	49

on psychoanalytic training by the School's interviewers with the later status of the applicants, but the only information available on later status is whether or not the man actually became a candidate. Ideally, one would want to know his *performance* as a psychoanalytic candidate. Dr. A made predictions on 97 applicants to the Menninger School of Psychiatry. A high percentage (47 per cent) of the 30 he recommended to "Reject" for psychoanalytic training had by the time of the follow-up dropped out of psychiatry completely. Only 1 of the 27 he said to "Take" for psychoanalytic training dropped out of psychiatry. Furthermore, most (63 per cent) of those he recommended to "Take" for psychoanalytic training became candidates. Only 27 per cent of those he voted to "Reject" became candidates anyway. In summary, the crude, three-category judgments made at an early stage of a man's specialization predicted moderately well his entering or not entering the field.

#### RECOMMENDATIONS

In light of the above considerations, we believe that the recommendations given in Chap. 17 apply as well, in general, to psychoanalysis as they do to psychiatry. When allowances are made for the different selection ratios, slightly differing emphases in personality requisites for the two types of work, and the more highly trained and preselected population of applicants for psychoanalytic training, a

great deal of what we have reported throughout this book can be helpful in psychoanalytic selection.

In this connection, recommendation No. 10 in Chap. 17 should be particularly underlined. Candidates for psychoanalytic training will almost invariably have had some training and experience in psychotherapy, and careful inquiry of supervisors who have given this training can furnish what is perhaps the most directly relevant information on a man's potentialities as a psychoanalyst. Of course, the more the type of psychotherapy a man has been doing differs from psychoanalytic technique, the more dubious is the transfer (see Chap. 15).

It may seem that if a man has completed part or all of a psychiatric residency, successfully dealing with disturbed patients and withstanding the pressures and rigors of his work as a resident, there is little need for psychological testing. The experience at the Topeka Institute for Psychoanalysis would argue against this view. Even men who had been originally tested and accepted into the Menninger School of Psychiatry, and who had performed well in the residency training program, occasionally showed evidences on later retesting of hidden but developing psychopathology of very serious degree. In a couple of instances very dramatic proof of the validity of these indications appeared in transitory psychotic episodes, even though the interviewing training analysts had not been convinced that the applicants in question were seriously disturbed. Moreover, as we have pointed out elsewhere in this chapter, psychological tests are potentially the best available techniques for assessing certain personality variables particularly relevant to psychoanalytic work.

Of all the recommendations we have considered making for psychoanalytic selection, the one about which we feel most conviction is that a psychoanalytic institute should set up procedures for conducting research on its own selection process. Such an undertaking need not be ambitious or highly refined methodologically in order to be useful. A minimal program might be described in terms of the following set of directives:

First, continue to scrutinize applicants in the way that is currently employed. Have every person who participates in the process of assessment make definite ratings. This will include the members of the training committee who interview the candidate; the psychological tester (if any); and those who evaluate application forms, reports from residency training centers, and similar documentary

material. These ratings do not need to be numerical; Take—Doubtful—Reject is adequate. At any other point in the process where evaluation and decisions are made, write them down. Arrange to have a secretary collect all such ratings, making up a folder for each candidate; she should make sure that the standard set of predictions are made for each applicant, plague people until they hand in their ratings, and record carefully who gives what rating to whom and when. If interviewers can be persuaded to put into writing what they think are the bases for their ratings, so much the better (*cf.* Chap. 14).

Then set up a similar system for collecting criterion data. If a candidate is advised to withdraw from psychoanalytic training, record the event, the date, and any other relevant information. Record the completion of each preparatory analysis, with the date, the total number of psychoanalytic hours involved, and (perhaps) the training analyst's prediction about what kind of analyst the candidate will make. Similarly, record progress through to completion of controls, getting ratings from each control analyst of the candidate's proficiency at the time the two of them finished working together on a case. In these evaluations take into account and describe the candidate's suitability for treating different types of patients and for different psychoanalytic functions (research, teaching, therapy, theoretical work). These are, of course, only suggestions about criteria; each institute will want to decide what measures it wishes to use. We have assumed that the minimum goal, at least, is to select people who will complete psychoanalytic training without untoward difficulties. Make a final decision on just how these criterion data will be combined into some relatively simple categorization of candidates (such as Superior, Adequate, Borderline, and Inadequate). Take this step before beginning the research proper, and inform each interviewer and predictor about the exact nature of the criterion he is trying to predict.

After a few years have elapsed, it will be possible to begin checking back and comparing the original predictive evaluation with the criterion data that have been accumulated. This process will show up which persons have been able to do the best job of selection, and a careful study of successes and failures will enable improvements in the process to be made. Statistical treatment of the data can be confined to simple techniques such as Chi-square, the results being displayed in charts similar to Tables 9.6 and 12.7.

Any such undertaking will take several years to accumulate use-

ful information, even though it need not take much of anyone's time at any particular point. It is mainly a matter of trying to do things in an orderly way, making sure to record and keep information that will be available in the natural process of selecting candidates, and then looking back to find out what predictors yielded the best results. If several institutes followed this sort of plan, they could, after a time, compare notes and learn a great deal from each other's successes and failures. Against the trouble and effort that would be involved in such undertakings can be balanced the professional obligation of any educational group to evaluate the job that is being done. Candidates and training analysts alike are spending many years of their lives in an expensive process; it seems hardly too much to ask that every effort be made to waste as little of this time as possible.

## OVERVIEW

IN THE PREVIOUS CHAPTERS WE OFFERED RECOMMENDATIONS TO PEOPLE who are selecting residents or who are doing research; what remains to be communicated is a set of reflections about the broader implications of this project.

### THEORETICAL IMPLICATIONS

In Chap. 7, and from time to time elsewhere, we have referred to the main theoretical issue that has grown out of the long series of researches forming our project's ancestry: the relative merits of the clinical *vs.* the actuarial or statistical approach to prediction. Even though our project was not designed in order to make a clear-cut test of one method against the other, our methods and findings have considerable bearing on the controversy. We have already stated our approach to the issues (Chap. 7) with a proposal for compromise: developing a method that will combine the best of both—the richness of clinical appraisal and the power of statistical-actuarial techniques of research. In our manuals we tried to carry out this program. We kept the clinician's tools, such as the TAT and interview, but sought to make orderly and explicit the ways they were to be used for this particular kind of prediction, employing as much of actuarial

procedure as we could (preliminary study of criterion groups, cross-validation, etc.). The validities we got in the final Predictive Study contain a good deal of promise for the method but are far from being a fair measure of its potentialities. Considering the small numbers of cues and the lack of reliability in both predictor and criteria, the validities achieved with the TAT Formal, the Self-Interpretation, and the Interview manuals are impressive. We hope that others will apply this approach to new predictive problems under more favorable circumstances.

In our attempt to apply the classical actuarial technique of multiple correlation, using relatively objective test scores as predictors, the cross-validation showed that this method of using the Strong, Rorschach, and Wechsler-Bellevue tests did not retain useful validity, whereas clinical evaluation of a battery in which the latter two were the principal instruments yielded predictions that had significant validity (Chap. 13).

We are particularly concerned with repeating and highlighting these points because our work has already been cited, from preliminary reports, as indicating the failure of the clinical method of prediction (*cf.* Cronbach, 1956). It is true that on the average the clinicians who made free predictive ratings in both Designs I and II did not achieve remarkably high levels of accuracy, not much better than the clinicians of the Michigan project who attempted to predict performance in clinical psychology (Kelly and Fiske, 1951). Two of our findings stand out in sharp contrast to those of Kelly and Fiske, however:

1. *Some individual clinicians achieved validity coefficients from about .5 to over .6, or about as good as those in the best reported actuarial attempts to predict success in vocational training.*

2. *No objective test predictor approached this level of validity in our study, whereas in the Michigan project the best results were achieved with certain test scores.*

It is true that one of our best predictors was the Verbal IQ from the Wechsler-Bellevue scale, but it was surpassed by several clinical predictions and failed to yield a cutting-score that would be as useful in selection as the clinical synthesis of the tests. And the Strong Vocational Interest Blank, which (with an intelligence test) gave the best validities in the Kelly-Fiske study, failed to yield any good

predictor of competence in psychiatry. Even the special key (Psychiatrist A), produced by Strong from a statistical analysis of thousands of blanks filled out by diplomates in psychiatry, failed to predict any of our criteria at a statistically significant level. This last finding deserves emphasis, because Strong's key was the product of a highly developed statistical technology, had an adequate numerical base and every opportunity to show what the actuarial method could achieve. As compared to our manuals for analyzing projective techniques, it had only the advantages of being completely objective and cheap to use; but despite their tiny numerical bases and their own crudities, the manuals showed that clinical judgment could be kept as a main ingredient in the predictive process and lead to significant validities.

If we had concentrated on an actuarial rather than a clinical approach and had come up with a simple, objective procedure that had a high and stable level of validity in predicting psychiatric performance, we believe that it could have been misused. It might have tempted many psychiatric training centers to adopt a single mold from which would have been cast a generation of psychiatrists who would have had to meet the problems of the future with a standard set of resources derived from the past. The more successful we are in finding objective, impersonal, and statistical methods of selecting members of a profession in the image of its past leaders, the more rigid will be the pattern into which it is frozen.

For a concrete example, consider Strong's Psychiatrist A key again for a moment. It expresses the pattern of interests held in common by men who were diplomates in psychiatry at the end of the war, most of whom must have been trained fifteen to twenty years ago. It should hardly be surprising that residents whose interests most closely approached this pattern tended to have skills as administrators and diagnosticians rather than as psychotherapists. If the key had happened to achieve a high correlation with our over-all criterion, it might have helped populate American psychiatry of the 1960's and 1970's with near-replicas of the old state-hospital superintendent.

It might be argued, however, that a similar result could have been expected if we had succeeded in providing explicit methods of clinically analyzing other types of data to select psychiatric residents. They, too, would have been based on a study of men who were successful at one time in history and would have suffered the same danger of becoming out-dated. The answer is that even sophisticated clinical

prediction never gets quite that rigid. Changes creep in; the result may be that validities gradually regress, or the drift may be determined by valid appraisals of newly important variables. Clinical methods are more flexible than their actuarial counterparts: They can be more readily modified by new studies based on observations of developing trends in the criterion. Moreover, valid clinical impressions can be obtained from an intensive study of a few known cases, whereas it takes large samples to set up or revise an actuarial system. There can be no guarantee that clinical methods *will* be kept up to date, of course, nor that the attempt to do so will not spoil their validities. Any predictive system needs constant overhaul and revalidation.

By staying with the capriciously accurate, sporadically reliable, and eminently flexible method of clinical judgment in selecting trainees, psychiatry will at least be able to keep in touch with developments in a growing and changing profession. Moreover, it will be able to maintain a healthy diversity within its ranks. There are many jobs to be done in psychiatry, requiring quite different kinds of men. There must be thoughtful persons who like to sit in deep chairs and analyze patients all day long. There must be activists to organize new institutions and give inspirational leadership to groups of colleagues. Psychiatry needs many more men and women than it has whose main interest is in research and teaching; others to work with broad preventive programs in public health; group therapists; specialists in somatic treatments; and many more varieties of the general species. If the purely actuarial approach were to be seriously applied to psychiatry, it would be necessary to develop a formula for each of many different types of practice and to revise it constantly as new developments created needs for new types of practitioners. To do so would be impossibly expensive and laborious. Psychiatry is well-off, therefore, adhering to a basically clinical approach to assessment and prediction in selecting its members, but trying constantly to make it more scientific.

The important issue, however, is not which method of selection is best for any particular profession but the relative inertia of actuarial predictive systems and the maneuverability introduced when the generating of predictions is done by clinical judgment. This freedom is a source of weakness as well as strength: it enables the clinician to fall into errors of many kinds (to which statistical predictions are less subject) and also to adapt himself sensitively to all kinds

of changing circumstances. When clinical methods are given a chance —when skilled clinicians use methods with which they are familiar, predicting a performance about which they know something—and especially when the clinician has a rich body of data and has made the fullest use of the systematic procedures developed by actuarial workers, including a prior study of the bearing of the predictive data on the criterion performance, then sophisticated clinical prediction can achieve quite respectable successes. We hope that clinicians will take some heart from our results, but urge them to refine their procedures by learning as much as possible about statistical prediction and adapting it to their own ends.

#### PRACTICAL IMPLICATIONS

In considering the practical implications of any research findings the first question to ask is: how far may they be generalized? In many ways it was an advantage to have conducted the study entirely within one community and one school, but we cannot avoid facing the question: How directly can the results of this study be used in Minnesota, California, or North Carolina? Each institution ought to answer this question for itself by the sort of research we have recommended, yet a general answer is also possible.

The heart of the matter, as we see it, is the representativeness of the Menninger School of Psychiatry as a training center, and the typicality of its program.

Fortunately, a recent nationally-based survey of residents in fourteen major institutions throws a good deal of light on this problem (Ward, 1954). The psychoanalytic orientation of the Menninger School of Psychiatry is becoming more characteristic of psychiatric residencies as time goes by. Ward reports: "The orientation of most of the centers is described as Freudian or Neo-Freudian; however, the resident's own orientation is apt to be in this direction regardless of his center's orientation." The large majority of residents reported an interest in psychoanalytic training or treatment, just as the residents have consistently done in the Menninger School of Psychiatry. A report by the Committee on Medical Education of the Group for the Advancement of Psychiatry contains findings of surveys of most of the psychiatric training programs in the country, with very similar conclusions (Greenhill *et al.*, 1955). Even the size of the Menninger School, the extensiveness and formalization of its didactic program,

no longer distinguish it so sharply from those of other training institutions; over the ten years since its beginning in 1946, there has been a process of convergence. Other centers have increased in size as the Menninger School of Psychiatry, by decentralizing its residents over several institutions, has in effect grown somewhat smaller; and there is a growing resemblance between the curriculum of the Menninger School of Psychiatry and those of the other major psychiatric training institutions.

For these reasons, therefore, the qualities that make an applicant a good resident in the Menninger School of Psychiatry are probably very much the same as those necessary for success in most other psychiatric residencies; and the same goes for the negative indications.

In discussing generalizability, however, we want to call attention again to the variety of possible and actual roles in psychiatry mentioned just above and to the fact that the Menninger-trained psychiatrist typically aspires to be a psychoanalytically-oriented psychotherapist. The evidence argues that this is to be the predominant pattern for the immediate future in American psychiatry, but it is *not* put forward as the only useful and needed one.

The method of gearing our selection techniques not directly to job descriptions but to *intervening personality variables* makes the findings potentially much more widely applicable. For example, we did not ask simply: "Can this man become a psychiatrist?" but: "Does he have the qualities of sensitivity, intelligence, integrity, etc., that are relevant to his becoming a psychiatrist?" If a psychological analysis of the work done in another occupation suggests that some of the same aspects of personality are involved, it may be possible to apply some of our methods to rather different kinds of selection. Landsberger's success in using our TAT Formal manual to predict success in labor mediation is a practical demonstration of this possibility (Chap. 12). It would be unwise to apply our findings so far afield on an operational basis, however, without first establishing by research that the extension was a valid one.

Before inquiring further into implications of our findings, it will help to look for the *underlying problem*. The presenting problem that led to selection research was the many applicants for an expanded training program. This growth in numbers of both interested physicians and residencies came about not only because the

psychological health needs of the nation were increasing but particularly because people were becoming aware that many of their troubles were of a psychiatric nature, and that means of alleviating them existed. In Chap. 2 we reviewed the extent of demand, and the facilities for satisfying it, at the end of World War II; but ten years later the gap between needs and resources seems just as wide—it is still estimated that we need twice as many psychiatrists as we have right now (U.S. Congress, 1953), and prospects are for continuing great increases in demand for psychiatric and allied services as both the population and the rate of diagnosed mental disorder in the population continue to rise (Albee and Dickey, 1957).

Behind the presenting research problem—improving methods of selecting psychiatrists—lies a more basic problem, then: how are we to obtain enough psychiatrists of the highest caliber? *Improved selection* is only one possible solution. Two others are *intensified efforts at recruitment* and *improved methods of psychiatric education*.

We should not speak as if all that was necessary was to increase the numbers of good psychiatrists, for two reasons: First, persons trained in other professions (*e.g.*, psychiatric social work, clinical psychology, psychiatric nursing) play important parts in meeting mental health needs, and much of what we shall say about psychiatrists will apply equally to them. Second, in thinking about the future we should not be limited by a simple extrapolation of present trends; it is quite possible that the present division of labor between existing professions is not the best that could be achieved, and changing times may bring a reshuffling of functions and types of training in which the existing professions will alter and new types of specialists will emerge.\*

Improved methods of selecting and educating psychiatrists will enable us to make the best use of available resources and will improve the quality of the doctors trained, but only by active *recruitment* can the quantity of psychiatrists be greatly augmented. And considering the great shortages of today, the highest priority should be given to increasing the numbers of psychiatrists, selected and trained as well as possible. To do so will require more residencies at more psychiatric training centers, but since there are at present many psychiatric residencies going begging (Anderson *et al.*, 1956),

\* Kubie (1948, 1950) offers a bold suggestion for a new professional school and degree in which essential aspects of medical and psychological training will be combined.

increasing the numbers of physicians who want to specialize in psychiatry would seem to be the first order of business.

As the teaching of psychiatry in medical schools is improved, as more effective therapies in psychiatry become generally known, as the general spread of information about mental-health needs continues, the natural course of events may bring increased numbers of applicants. Moreover, the "natural course" can apparently be hastened; after several of the Menninger School's staff had visited the medical school at Saskatchewan for lectures and teaching, the following year a quarter of the new residents at Topeka State Hospital came from this single province (population less than 900,000) of Canada.

Here we come up against another barrier, however: psychiatry is not the only type of medical practice in which a shortage of personnel is felt. Moreover, other medical specialties are growing faster: "Despite strenuous efforts made in recent years to encourage more young physicians to enter into psychiatric residencies, there has been no proportionate increase and, in fact, there has been some ground lost in the past three years. . . . For many years the proportion of physicians indicating a limited specialty in psychiatry has remained quite constant between the limits of 5½ per cent to 7 per cent" (Albee and Dickey, 1947, p. 62). When one considers that half the hospital beds in the country are psychiatric and only about one physician in 20 is a certified psychiatrist (psychiatry has only 15 per cent of the specialists), it is obvious that psychiatry needs the lion's share of the new physicians for some time to come and is not likely to get them unless "strenuous efforts" are continued and expanded.

The shortage of physicians of all kinds and the poor prospects for increasing their numbers make the outlook for an early great increase in the number of psychiatrists even dimmer. It is difficult to decide just how many physicians per unit of population the country should have, but it is clear that the relative numbers of doctors have not risen during the past thirty-five years. Indeed, applications to medical schools have been falling off in recent years; according to Benjamin Fine in *The New York Times* of September 6, 1955, the ratio of applicants to medical-school openings had declined from 7-1 down to 2-1. In some states where only local citizens may apply to the state-supported medical colleges, schools have difficulty in filling their classes.

Yet there would undoubtedly be a great increase in qualified applicants to medical schools if there were no religious, racial, sex-

ual, geographic, or economic barriers to medical training. It seems likely that there are enough young people with the intellectual and personal qualifications, and the interest, to fill all the medical schools that exist and that have any chance of being built in the next few years.

Even without such far-reaching changes in policy, pressure of applicants on colleges of medicine is bound to increase with time, yet medical schools' capacity can be increased only quite slowly. If the American Medical Association were to drop its opposition to any form of federal aid to medical education, government funds might become available to underwrite the enormous cost of creating new medical schools. But there would still be the problem of finding faculties to staff these schools; today, good medical teachers are said to be harder to find than money for buildings.

It seems likely, therefore, that there will be a shortage of physicians for many years to come. Under the circumstances, perhaps the most that could be asked of medical schools is that they modify their selection policies slightly away from a heavy emphasis on scholarship in the natural sciences and toward seeking out the type of person we have described in these pages as a potentially good psychiatrist. If medical schools are to be equal to the task of training physicians who can practice the medicine of the late twentieth century, they must pay more attention to the human qualities of the applicant, as well as to the psychiatric part of the curriculum.

Just as psychiatry is not the only branch of medicine urgently in need of manpower, medicine is only one of the professions competing for the top 5 per cent in intellectual ability of each crop of young men and women. The reports of the Research Staff, National Manpower Council (1953) and the California Institute of Technology (Weir *et al.*, 1956) highlighted this crisis in the use of our most precious resource, human talent. They showed how far behind the Soviet Union we have fallen in the production of scientists, engineers, and physicians; how we need to double our production of such trained minds at once. Our national survival may depend on how well we solve this problem. Some kind of nationwide program of scholarships seems inevitable to enable bright high-school students to go on to college and postgraduate training in scientific research, medicine, or one of the other critical professions.

The task of recruitment, thus, is quickly seen to involve the necessity of social planning on a large scale. Ultimately, it may

require a program of conveying information about needed types of professionals to young people during their adolescent years, when vocational choices begin to be seriously considered, so that a potential psychiatrist will be sure to learn of all the opportunities that are available to him.

After these sobering thoughts on the difficulties of increasing the quantity of psychiatrists, let us consider how the quality may be maintained or even improved. During the past ten years we have seen so many highly intelligent, serious, honest and well-trained young physicians go into psychiatry that we are not worried about the issue of quality, except to maintain it as numbers increase. The general implications of our results may be read in either of two different ways.

On the one hand, one could argue as follows: Our results show that using one or two unselected interviewers, or a psychologist interpreting a couple of projective tests, is a waste of money—just as good results could be obtained from a careful study of credentials; but when carefully selected clinicians combine the results of intensive study of an applicant, a decidedly useful increase in selective efficiency comes about. This increase is great enough to warrant the necessary investment of the time and money when we consider the cost of training to the institution and to the resident, and the dangers to patient and resident alike of taking into training someone not suited to the field.

On the other hand, it can be maintained that early reporting of new residents' work and adjustment could turn up the hopeless misfits before anyone has made much of an investment in training them; moreover, the long-run results of our study show that many residents who floundered badly at first did become Late Developers, and even many members of our original Small Sample Lows have become respected and undoubtedly useful psychiatrists. Don't lose patience with the slow developers, this argument continues; remember that we need to keep in psychiatry every man we can get—let's make the most of each resident's potentialities. If a man has no talent for psychotherapy, perhaps he will make a good neuropsychiatrist working with somatic treatments, or a good social psychiatrist working in industry or public health, or a hospital administrator who will be familiar with the problems of the men who are treating patients even if he does not take part in it himself. Let us teach him to know and respect his limitations and his strengths. After all, there

is no legal barrier keeping a physician who has been rejected or dropped from training from setting up a psychiatric practice anyway, and thus barging into all kinds of situations he could be trained to avoid. How much better to find a niche for him and then teach him the value to all concerned of staying within the limits of his abilities!

We believe one can extract and use the truth in both these interpretations. After doing as good a job as possible with selection, and thereby ruling out at least some of the unfit applicants, our schools of psychiatry need to put stress upon training—about which we know too little. Educational research has been going on for decades, but there has been hardly any systematic and controlled study of the process of psychiatric training. The need is obvious enough to the people who are responsible for it. In a recent survey, the thoughtful discussions of training by many directors of psychiatric residency programs and of psychoanalytic institutes, facing frankly their ignorance of how to train best, expressed their conviction that much can be done to improve present teaching methods and mentioned the difficulties of making progress (Alexander, 1956).

Research directed toward improving training methods in psychiatry will aim at several objectives. How to give the best training to the greatest number of residents with the smallest expenditure of staff time will probably be foremost, yet we should like to enter a few special pleas. The need for research in psychiatry by psychiatrists who know their own field and who at the same time have an adequate grasp of research method is so great, and the number of such research psychiatrists so small, that special attention will have to be given to finding ways to teach research thinking and research techniques to more men, starting while they are residents.

Inquiry should also be directed, we believe, especially toward the barriers to learning psychiatry experienced by the *least* promising students. Research in training will tend to improve teaching for all, and thus raise the general level of competence, but the talented ones will manage very well for themselves if attention is focused on the less obviously gifted. In the Menninger School of Psychiatry we have observed that many residents develop sensitivity, psychological-mindedness, empathy, and other such assets, although on initial assessment we failed to see much sign of them. Perhaps a special study of such men, attempting to find out how they were helped to grow in these ways, would provide leads that could help open psychiatry to

many sincerely interested, intelligent people who seem to lack special talents.

The role of personal treatment in the training of psychiatric residents needs to be studied more carefully. The amount and kind of therapeutic help required by different kinds of residents, its timing, the degree to which it is best recognized, recommended or ignored by supervisors—these and other facets should be examined. It may turn out that almost all psychiatrists undergo some kind of treatment at some point either during residency or in the years directly afterward (as seems to have been true of our research subjects). If such is the case, it might be wise to provide facilities through which treatment of the appropriate kind can be obtained by *all* residents, perhaps after the pattern of psychoanalytic training. Such a policy would require re-examining several problems of selection and education.

Directing future research toward training would have a number of indirect benefits. It might lessen the atmosphere of competition and make the residents feel that their teachers were concerned with learning to do their own tasks better. After a man is selected it is consistent with a respect for human individuality to try to find ways of helping him learn and grow so that he can fill a useful role in the profession of his choice.

We have considered the possible consequences of concentrating on selection, recruitment, or training as approaches to the basic problem of satisfying mental-health needs. The experience of carrying out our research, as well as some of its findings, convinces us that it is education that will in the long run repay increased cultivation with the richest yields. Probably it is not necessary to add that all three fields must be tilled, and that all three can produce much of value. We like particularly the thought of a new kind of selection research closely geared in with an experimental approach to education: In such a project, an applicant would be carefully studied, and his strengths and weaknesses diagnosed, not just to accept or reject him, but in order to plan the kind of program that he needs most to develop with least conflict and difficulty. Perhaps this is a Utopian vision of a more individualized kind of education than will be possible in psychiatry for a long time to come. But the aim of learning ways to bring out the best qualities people have latent within them, whether we do so as scientists, religionists, therapists, or educators, is a useful and satisfying orienting ideal.

## STATISTICAL NOTE

(For Readers Who Know Little about Statistics)

SEVERAL STATISTICAL TECHNIQUES ARE USED IN THIS RESEARCH AND REFERRED to in this volume; some of them may not be familiar to all readers. But if the reader will scan this note and then refer back to it from time to time when he runs across references to particular statistical methods, it should be possible for him to follow and understand the entire presentation and discussion.

### CORRELATION

Let us start with correlation, which is measured by the correlation coefficient,  $r$ .\* A correlation coefficient is a single statistical measure that provides a way of expressing quantitatively the degree of association between two variables. Suppose you have two sets of numbers; for example, predictive ratings and criterion ratings, based on the same group of subjects. If you line them up in two columns, pairing the two scores for each subject, and inspect them, you may find that there is a tendency for the people who get high scores in one column to get high scores in the other, whereas those who get low scores on the predictor get low scores on the criterion. This is what we refer to as association or correlatedness; a correlation coefficient gives us a simple and easily computed measure of it.

Correlation coefficients range from +1.00 through 0 to -1.00. Off-hand, one might think that a complete lack of association between two sets of numbers might give the "lowest" correlation, -1. This is *not* the case. The lowest correlation is .00; this represents a condition in which the numbers in the two columns are randomly arranged with respect to each other. But suppose, when we look over the columns, we find that a high number in one column tends to go regularly with a low number for the same man in the other column. This is the state of affairs known

\* There are several measures of correlation, of which  $r$ —the Pearson product moment correlation coefficient—is only one, though it is the most used. All correlations in this book are of the Pearson type.

as *negative correlation*. If we have a high negative correlation (up to  $-1.00$ ), we can confidently predict that anyone who gets a low score on a predictor will get a high score on a criterion measure, and if this holds up in repeated groups, then we have useful information. For all practical purposes, however, we can neglect negative correlations in considering the results of this project, for, with a very few exceptions, the coefficients are positive, and negative ones are almost all so low that they are practically equivalent to  $.00$ .

Suppose we are considering a particular correlation,  $r = .50$  between a predictive score and a criterion measure. What does it mean? In general, it means that there is a tendency for high-scoring subjects on the predictor to turn out high on the criterion, and for low-scorers to get low criterion scores. But that is just restating the meaning of a correlation; it tells us little about how *good* this one is.

There are two ways of evaluating the "goodness" of a particular correlation coefficient. One is to assign some kind of qualitative meaning to correlations of various sizes; the other is to ask how great is the likelihood that the correlation would hold up if the study were done again (which is the issue of "significance").

Let us consider the first approach. Regardless of how significant it is, or of how many subjects it is based on, a correlation of  $.50$  always indicates the same amount of predictive power, and a usefully large amount it is. It will probably be helpful to think of the correlations discussed in this book in terms of the following very rough qualitative equivalents:

- .00 to .19: No correlation or a negligible amount.
- .20 to .29: Slight, barely usable correlation.
- .30 to .39: Modest but useful correlation.
- .40 to .49: Fairly good to good.
- .50 to .59: Good, definitely helpful.
- .60 and over: Very good; correlations much over  $.6$  are quite rare in most psychological prediction.

We should add that the above qualitative scale might not apply to correlations used in another sort of problem. A correlation of  $.6$  is still quite far from enabling one to predict the correlated score with high accuracy, but a predictor that was that good would be very useful to anyone predicting success in school, and probably the best one available.

Now for the issue of *significance*. This is an important concept in statistical reasoning, and we shall have reference to tests of the significance of several kinds of statistics. Let us start with a hypothetical example: Suppose we were interested in some easily visualized matters such as height and weight. We take 10 men, weigh and measure them, and correlate the resulting figures. Now we know from general observation that tall men tend to be heavier than short ones, and controlled studies have found consistent, low positive correlations between these measures. But there are lots of exceptions: tall spare fellows and short heavy ones.

With a sample as small as 10, we run the risk of picking up a few of these unusual cases, thereby distorting the results. The larger the sample, the less an atypical subject can affect the correlation. Consequently, if we take 20 random samples of 10 men each, we can generally expect to get one correlation as high as .65 purely by chance, even if the true correlation in the population being sampled was zero. The correlation coefficients we should derive from 20 samples of 100 men would vary much less, and we could be more certain, therefore, that the figure we got from any one of them was close to the true correlation in the population being sampled.

Consequently, the size of the sample determines how significant a correlation is. When we say a correlation is significant, we mean it is significantly (or reliably) different from zero at a certain level of confidence. The level of confidence (and thus significance) is usually expressed in terms of the so-called *p-value*—meaning probability of occurring by chance; for example,  $p = .05$ . In the example above, of the correlations based on 10 cases, the statement that an  $r$  of .65 could occur once in 20 tries with such small samples, even if the variables were unrelated in the total population, has the same meaning as the statement that the correlation of .65 is significant at the 5 per cent level, or that  $p = .05$ . If a result is likely to occur as seldom as once in 20 by chance, we accept it as probably not due to chance and thus agree to tolerate that many mistaken decisions. Notice that a correlation of .50 (which we have described as “good, definitely helpful,” in terms of its absolute magnitude) could occur quite frequently by chance with samples of only 10 subjects; so if we found that  $r = .50$  with a sample of 10, we should not take it seriously—it would be “insignificant” even though large! In practice, of course, if one got such an encouragingly large but insignificant correlation with a small sample, one would increase the size of the sample and see whether it was, in fact, a fluke or whether it held up and finally reached statistical significance.

Since significance levels for correlation coefficients depend on sample size, a correlation of .25 may be quite unworthy of consideration in one study and in another may represent a highly significant finding. We should never forget that these two aspects of a correlation are relatively independent—a correlation of .10 is significant at better than the .05 level if based on a sample of 1200 subjects; but its predictive power, and thus its practical usefulness, remain negligible.

#### TWO-TAILED AND ONE-TAILED TESTS

Now how do we determine the significance of any particular correlation? Certain formulas derived from probability theory tell us just how many correlations of any size could be obtained from a great number of samples of a given size, if the true correlation in the population being randomly sampled is zero. (This hypothetical state of no relationship is called the “null hypothesis”—it is what we want to be able to reject with

a certain degree of confidence.) But notice that it can be rejected in either of two ways: the two variables might be highly correlated positively or negatively. Half of the time, in sampling from a zero-correlation population, sample correlations will be positive, half the time they will be negative, so we have to consider both extremes of the hypothetical distribution of possible sample findings. These extremes (the few large correlations that occur by chance) are the "tails" of the distribution. Thus, when we want to be able to say with confidence (willing, for example, to be wrong once in 20 times—the .05 level) that it is *not* true that the variables are uncorrelated, we have to count the positive values that could be obtained by chance, and the negative ones as well. So, for the 5 per cent level, we take the most extreme  $2\frac{1}{2}$  per cent at each tail of the curve.

Suppose, however, we are interested in rejecting another kind of null hypothesis, as we are in most places throughout this book: we want to disprove the statement that predictors and criteria are not *positively* correlated. If an interviewer's judgments about a candidate are not positively correlated with a criterion of subsequent performance, we don't care whether the actual correlation is zero or negative—in either event we can't use him. Therefore, in such instances as this, it is necessary to consider only one tail of the theoretical distribution—those high *positive* correlations that could come about 5 per cent of the time from random-sampling variations alone. Under these "*one-tail*" rules, somewhat smaller correlations will meet the .05 standard of significance, which we call the 5 per cent *point* instead of *level* to focus attention on the fact that we are dealing with a restricted (directional) null hypothesis.

#### SIGNIFICANCE OF MEAN DIFFERENCES

These general principles about significance and p-values apply to other statistics besides correlations. If we are comparing two samples on some measure (*e.g.*, the IQ's of college students and those of psychiatric residents), we will almost always find that the average IQ in one group is higher than that in the other and also that the two distributions overlap—the brightest man in the group with the lower mean has a higher IQ than the dullest man in the other group. Even if we take different samples from the same population, this same state of affairs is likely to exist. We need some way of establishing, with some specified degree of assurance, that any obtained difference is not just due to sampling fluctuations. Once again, therefore, we set up a null hypothesis and try to reject it: either that there is no difference between the samples (because they come from the same population, which requires a two-tailed test of significance, since it could be rejected if the students' IQ's are higher than the residents, or vice versa); or that the more educated group does not have a higher average IQ than the less educated, which requires a one-tailed test since only one kind of chance difference is relevant to it. The test used in this book is Student's *t*, which gives either a one-tailed or two-tailed p-value for any mean difference (*i.e.*, a statement of the

percentage of the time you would get differences of the obtained size from pairs of samples randomly drawn from the *same* population).

### CHI-SQUARE

We sometimes refer also to a statistic called Chi-square, another way of testing the difference between two sets of numbers, which may sometimes be interpreted as an association between two variables. Consider the kind of results discussed in Chaps. 9 and 12, when the predictor is a judgment of Take or Reject, and the criterion is a pass-fail one such as Satisfactory or Unsatisfactory. All possible cases can be classified in the following table:

CRITERION	PREDICTOR	
	Take	Reject
Satisfactory	30	5
Unsatisfactory	7	10

There are several techniques for enabling us to find out how often such a distribution of data might occur by chance (which still means by random-sampling fluctuations); Chi-square is one. The other used in this research is Fisher's exact test, which is needed when there are small numbers of subjects. By either test, the above arbitrary set of "results" would be significant at better than the .01 point (one-tailed test). They may be interpreted to mean either that candidates about whom the interviewer said "Take" are more often Satisfactory than those he wanted to Reject or that there is a significant relation between the predictive judgment and the criterion.

### MULTIPLE CORRELATION

In Chap. 13 there occur a few references to a method of statistically combining several predictors and bringing them to bear on a single criterion. The logic of multiple correlation is not difficult, even though the formulas are complicated. Suppose you were trying to predict college grades of entering freshmen. You have their average grades in high school and the results of an intelligence test. With earlier samples of students, each of these predictors perhaps correlates .4 with the college-grades criterion. Surely we can do better than that with both of them. Just how much better? The multiple correlation technique tells us. The important question it takes into consideration is: how highly correlated are the two predictors with each other? Obviously, if IQ and high-school grades were almost perfectly correlated with each other ( $r = .99$ ), the information contained in one set of scores would be merely duplicated by the other, so it would add virtually nothing. The more independent of each other two predictors are, the better multiple correlation coefficient you get when you combine them. The formula gives the weighting of each that will result in the largest final correlation.

**FACTOR ANALYSIS**

IQ and high school grades are not, in fact, correlated nearly as highly as .99 (.4 would be more like it), but they are significantly related to each other as well as to college grades. If we were to give tests of reasoning ability, reading comprehension, and memory for meaningful material to the same sample of entering freshmen, the chances are that these tests would all correlate with college grades, with the other two predictors, and with each other. If we then set up a table displaying the correlation of each of these measures with one another, we might find appreciable coefficients in each cell. This state of affairs suggests that all of the tests and the two kinds of grades are measuring slightly differing aspects of the same thing, which in this case we should probably take to be general intellectual ability.

Factor analysis is a way of determining the minimum number of mathematical entities (called factors) that need to be assumed to account for any given particular pattern of correlations. It also tells just how much contribution to a factor is made by each of the variables. In our example it seems reasonable to expect that the intelligence test would be the best single measure of the general factor that causes all the measures to be intercorrelated, and thus would have the highest *loading* on that factor. After the first factor was extracted, it might turn out that other factors would appear that were not apparent because of the predominant effect of the first one, such as, perhaps, a specifically verbal ability. Or it might turn out that the general factor accounted for all of the relationships between the variables, leaving only small residual variations unique to each test or set of grades. When we can find psychological constructs to link up with the mathematical factors, a factor analysis often uncovers an underlying order in a confused situation.

## BIBLIOGRAPHIC REFERENCES

- ABEL, THEODORA, OPPENHEIM, SADI, & SAGER, CLIFFORD J. (1956): Screening applicants for training in psychoanalytically oriented psychotherapy. *Amer. J. Psychother.*, 10, 24-39.
- ADORNO, T. W., FRENKEL-BRUNSWIK, ELSE, LEVINSON, D. J. & SANFORD, R. N. (1950) : *The authoritarian personality*. New York: Harper.
- ALBEE, GEORGE W. & DICKEY, MARGUERITE (1957): Manpower trends in three mental health professions. *Amer. Psychologist*, 12, 57-70.
- ALEXANDER, FRANZ (1956) : *Psychoanalysis and psychotherapy; developments in theory, technique and training*. New York: Norton.
- ALLPORT, GORDON W. (1937) : *Personality, a psychological interpretation*. New York: Holt.
- AMERICAN PSYCHIATRIC ASSOCIATION (1953) : *Psychiatric training in the United States and Canada. A descriptive directory*. Washington, D.C.: the Association.
- AMERICAN PSYCHOANALYTIC ASSOCIATION (1956) : Minimum standards for the training of physicians in psychoanalysis. *J. Amer. psychoanal. Ass.*, 4, 714-721.
- ANDERSON, DONALD G., MANLOVE, FRANCIS R., & TIFNER, ANNE (1952) : Medical education in the United States and Canada. Fifty-second annual report of medical education. *J. Amer. med. Ass.*, 150, 99-148.
- ANON. (1949) : *Dictionary of occupational titles*. Vol. I, 2nd ed. Washington, D.C.: Government Printing Office.
- BALINT, MICHAEL (1948) : On the psycho-analytic training system. *Int. J. Psycho-Anal.*, 29, 163-173.
- BARRON, FRANCIS X. (1950) : Psychotherapy as a special case of personal interaction: prediction of its course. Unpublished doctoral dissertation, Univer. of California.
- BARRON, FRANK (1953) : An ego-strength scale which predicts response to psychotherapy. *J. consult. Psychol.*, 17, 327-333.
- BARRON, FRANK (1954) : *Personal soundness in university graduate students. An experimental study of young men in the sciences and professions*. Berkeley and Los Angeles: Univer. of California Press.

## 366 Bibliographic References

- BARTEMEIER, LEO H., KUBIE, LAWRENCE S., MENNINGER, KARL A., ROMANO, JOHN, & WHITEHORN, JOHN C. (1947) : Combat exhaustion. *J. nerv. ment. Dis.*, **104**, 489-525.
- BAVELAS, ALEX (1952) : Communication patterns in problem-solving groups. In Heinz van Foerster, Ed., *Cybernetics; circular causal and feedback mechanisms in biological and social systems*. New York: Josiah Macy, Jr. Foundation.
- BECK, SAMUEL J. (1944) : *Rorschach's test*. New York: Grune & Stratton.
- BERENGARTEN, SIDNEY (1949) : A pilot study to establish criteria for selection of students in social work. Paper read at Amer. Ass. Schools Soc. Work, Boston.
- BERENGARTEN, SIDNEY (1951) : A pioneer workshop in student selection. *Bull. N.Y. Sch. soc. Work*, **44**, (4), 3-12.
- BETTELHEIM, BRUNO (1947) : Self-interpretation of fantasy: the Thematic Apperception Test as an educational and therapeutic device. *Amer. J. Orthopsychiat.*, **37**, 80-100.
- BINGHAM, JUNE (1953) : *The inside story; psychiatry and everyday life*. Text written by June Bingham, with the collaboration of Jacob Levine. Compiled under the direction of Fritz Redlich. New York: Knopf.
- BINGHAM, WALTER V. (1939) : Halo, invalid and valid. *J. appl. Psychol.*, **23**, 221-228.
- BLAIN, DANIEL, et al. (1953) : Training hospital administrators. Sect. 7, pp. 117 ff. in *Steps forward in mental hospitals*. Proceedings of the Fourth Mental Hospital Institute, American Psychiatric Association, Mental Hospital Service.
- BLENKER, M. (1954) : Predictive factors in the initial interview in family casework. *Soc. Serv. Rev.*, **28**, 65-73.
- BLOOMGARDEN, LAWRENCE (1953) : Medical school quotas and national health. *Commentary*, **15**, 29-37.
- BRILL, A. A. (1949) : *Basic principles of psychoanalysis*. New York, Doubleday.
- BROSIN, HENRY W. (1948) : Psychiatry experiments with selection. *Soc. Serv. Rev.*, **22**, 461-468.
- BUNKER, HENRY A. (1944) : American psychiatry as a specialty. Pp. 479-506 in Hall, J. K. (1944).
- BURGESS, E. W., & COTTRELL, LEONARD S., JR. (1939) : *Predicting success or failure in marriage*. New York: Prentice-Hall.
- BURGESS, E. W., & WALLIN, PAUL (1953) : *Engagement and marriage*. Chicago: Lippincott.
- CHAMPNEY, HORACE, & MARSHALL, HELEN (1939) : Optimal refinement of the rating scale. *J. appl. Psychol.*, **23**, 323-331.
- Council of State Governments (1950) : *Mental health programs of the forty-eight states*. Chicago: the Council.
- CRONBACH, LEE J. (1952) : *Essentials of psychological testing*. New York: Harper.
- CRONBACH, LEE J. (1956) : Assessment of individual differences. Pp. 173-196 in Farnsworth & McNemar (1956).

- CRONBACH, LEE J., & GLESER, GOLDINE (1955) : Processes affecting scores on "understanding of others" and "assumed similarity." *Psychol. Bull.*, **52**, 177-193.
- CURETON, EDWARD E. (1957) : Recipe for a cookbook. *Psychol. Bull.*, **54**, 494-497.
- DAILEY, J. T., BROKAW, L. D., & LECZNAR, W. B. (1948) : *The development of the Airman Classification Test Battery*. Research Bulletin 48-4. Barksdale AFB: Headquarters, Air Training Command.
- DEEMER, WALTER L., & RAFFERTY, JAMES A. (1949) : Experimental evaluation of the psychiatric interview for prediction of success in pilot training. *J. Aviat. Med.*, **20**, 238-250.
- DERI, SUSAN (1949) : *Introduction to the Szondi test*. New York: Grune & Stratton.
- DEUTSCH, ALBERT (1937) : *The mentally ill in America: a history of their care and treatment from colonial times*. Garden City, N. Y.: Doubleday, Doran.
- DEUTSCH, ALBERT (1944) : The history of mental hygiene. Pp. 325-366 in Hall, J. K. (1944).
- DEUTSCH, ALBERT (1948) : *The shame of the states*. New York: Prentice-Hall.
- DU BOIS, PHILLIP H., Ed. (1947) : *The classification program*. AAF Aviation Psychology Program Research Reports, No. 2. Washington, D.C.: U.S. Government Printing Office.
- DYMOND, R. F. (1947) : The relation of insight and empathy. *J. consult. Psychol.*, **12**, 228-233.
- EBAUGH, FRANKLIN G., & RYMER, CHARLES A. (1942) : *Psychiatry in medical education*. New York: The Commonwealth Fund.
- EBEL, ROBERT L. (1951) : Estimation of the reliability of ratings. *Psychometrika*, **16**, 407-423.
- EHRENREICH, GERALD (1955) : A brief report on psychological testing experience in the selection of candidates for psychoanalytic training. Unpublished paper presented to the Topeka Psychoanalytic Society.
- EKSTEIN, RUDOLF (1950) : Trial analysis in the therapeutic process. *Psychoanal. Quart.*, **19**, 52-63.
- EKSTEIN, RUDOLF (1953) : On current trends in psychoanalytic training. In Robert Lindner, Ed., *Explorations in psychoanalysis*. New York: Julian Press.
- ELLIS, ALBERT (1955a) : Psychotherapy techniques for use with psychotics. *Amer. J. Psychother.*, **9**, 452-476.
- ELLIS, ALBERT (1955b) : New approaches to psychotherapy techniques. *J. Clin. Psychol.*, (Monogr. Suppl.), **11**, 208-260.
- ENGLISH, HORACE B., & NAGY, GEORGE (1950) : Professional school admissions. In Douglas H. Fryer & Edwin R. Henry, Eds., *Handbook of applied psychology*. New York: Rinehart. 2 vols.
- ERIKSON, ERIK H. (1950) : *Childhood and society*. New York: Norton.
- ERIKSON, ERIK H. (1953) : On the sense of inner identity. In *Health and human relations*. New York: Blakiston.

## 368 Bibliographic References

- ERIKSON, ERIK H. (1956) : The problem of ego identity. *J. Amer. psychoanal. Ass.*, **4**, 56-121.
- ERON, LEONARD D. (1955) : Effect of medical education on medical students' attitudes. *J. med. Educ.*, **30**, 559-566.
- ESTES, STANLEY G. (1946) : Deviations of Wechsler-Bellevue subtest scores from vocabulary level in superior adults. *J. abnorm. soc. Psychol.*, **41**, 226-228.
- FARNSWORTH, PAUL R., & McNEMAR, QUINN, Eds. (1956) : *Annual review of psychology*. Vol. 7. Stanford, Calif.: Annual Reviews.
- FENICHEL, OTTO (1941) : *Problems of psychoanalytic technique*. New York: Psychoanal. Quart.
- FENICHEL, OTTO (1945) : *The psychoanalytic theory of neurosis*. New York: Norton.
- FINESINGER, J. E., et al. (1948) : *An investigation of prediction of success in naval flight training*. Washington, D.C.: Civil Aeronautics Admin., Div. of Res., Rep. No. 81. (Mimeo.)
- FLANAGAN, JOHN C. (1954) : The critical incident technique. *Psychol. Bull.*, **51**, 327-358.
- FLIESS, ROBERT (1942) : Metapsychology of the analyst. *Psychoanal. Quart.*, **11**, 211-227.
- FLIESS, ROBERT (1953) : Countertransference and counteridentification. *J. Amer. psychoanal. Ass.*, **1**, 268-284.
- FRANKLE, ALLAN H. (1951) : Tests in the selection of social workers. Paper read in symposium on selection of professional personnel, at Amer. Psychol. Ass., Chicago, Sept.
- FRANKLE, ALLAN H. (1952) : Rorschach human movement and human content responses in relation to field work performance in social case work training. Paper read at Amer. Psychol. Ass., Washington, D.C., Sept.
- FRENKEL-BRUNSWIK, ELSE (1948) : Dynamic and cognitive categorization of qualitative material. I. General problems and Thematic Apperception Test. II. Application to interviews with the ethnically prejudiced. *J. Psychol.*, **25**, 253-277.
- FRENKEL-BRUNSWIK, ELSE (1949) : Intolerance of ambiguity as an emotional and perceptual personality variable. *J. Pers.*, **18**, 108-143.
- FREUD, SIGMUND (1910) : Observations on "wild" psychoanalysis. Pp. 297-311 in *Collected papers*. Vol. II. London: Hogarth Press, 1928.
- FREUD, SIGMUND (1926) : *The problem of lay analyses*. New York: Brentano, 1927.
- FREUD, SIGMUND (1927) : Postscript to a discussion on lay analysis. Pp. 205-214 in *Collected papers*. Vol. V. London: Hogarth Press, 1953.
- FROMM-REICHMANN, FRIEDA (1950) : *Principles of intensive psychotherapy*. Chicago: Univer. of Chicago Press.
- GIEDT, F. HAROLD (1955) : Comparison of visual, content, and auditory cues in interviewing. *J. consult. Psychol.*, **19**, 407-416.
- GILDEA, MARGARET C. L., & GILDEA, EDWIN F. (1944-45) : Personalities of American psychotherapists: Mitchell, Salmon, Riggs. *Amer. J. Psychiat.*, **101**, 460-467.

- GINSBURG, SOL W. (1950) : Values and the psychiatrist. *Amer. J. Ortho-psychiat.*, 20, 466-478.
- GINZBERG, ELI, et al. (1951) : *Occupational choice*. New York: Columbia Univer. Press.
- GITELSON, MAXWELL (1948) : Problems of psychoanalytic training. *Psycho-anal. Quart.*, 17, 198-211.
- GITELSON, MAXWELL (1954) : Therapeutic problems in the analysis of the "normal" candidate. *Int. J. Psycho-Anal.*, 35, 174-183.
- GLOVER, EDWARD (1939) : *Psycho-analysis, a handbook for medical practitioners and students of comparative psychology*. London: Staples Press.
- GOLDHAMER, HERBERT, & MARSHALL, ANDREW (1953) : *Psychosis and civilization: two studies in the frequency of mental disease*. Glencoe, Ill.: Free Press.
- GREENHILL, MAURICE H., et al. (1955) : *Trends and issues in psychiatric residency programs*. Report of the Committee on Medical Education of the Group for the Advancement of Psychiatry. Topeka, Kans.: Group for the Advancement of Psychiatry.
- GREENSON, RALPH (1954) : On empathy and sublimation. Paper read at Amer. Psychoanal. Ass., New York, Dec.
- GREGG, ALAN (1952) : Emergent ability. *V.A. Tech. Bull.*, Series 10, No. 83.
- GUILFORD, J. P. (1948) : Some lessons from aviation psychology. *Amer. Psychologist*, 3, 3-11.
- GULLIKSEN, HAROLD (1950) : *Theory of mental tests*. New York: Wiley.
- HALL, BERNARD (1949) : Unpublished study (summarized in Whitehorn, et al., 1953, pp. 65-66).
- HALL, BERNARD, GANGEMI, MARY, NORRIS, V. L., VAIL, VIVIENNE H., & SAWATSKY, GORDON (1952) : *Psychiatric aide education*. New York: Grune & Stratton.
- HALL, J. K., et al., Eds. (1944) : *One hundred years of American psychiatry*. New York: Columbia Univer. Press.
- HAMILTON, SAMUEL W. (1944) : The history of American mental hospitals. Pp. 73-166 in Hall, J. K. (1944).
- HARTMANN, HEINZ (1955) : Notes on the theory of sublimation. *Psycho-anal. Stud. Child*, 10, 9-29.
- HEATH, S. ROY (1955) : Ways of life, a student development personality model. Unpublished mimeographed MS, summarized in Heath (1956).
- HEATH, S. ROY (1956) : Student "characters." *Princeton Alumni Weekly*, 56, (13), 8-11.
- HEIMANN, PAULA (1954) : Problems of the training analysis. *Int. J. Psycho-Anal.*, 35, 163-168.
- HENDRICK, IVES (1939) : *Facts and theories of psychoanalysis*. 2nd ed. New York: Knopf.
- HENDRICK, IVES, Ed. (1946) : Report of Conference on Post-War Problems of Psychoanalytic Training. New York: American Psychoanalytic Association. (Mimeo.)
- HERON, ALASTAIR (1954) : Industrial psychology. *Annual Rev. Psychol.*, 5, 203-228.

## 370 Bibliographic References

- HOLT, ROBERT R. (1950) : An approach to the validation of the Szondi test through a systematic study of unreliability. *J. proj. Tech.*, 14, 435-444.
- HOLT, ROBERT R. (1951) : The Thematic Apperception Test. Chap. 7 in H. Anderson, & G. Anderson, Eds., *An Introduction to projective techniques*. New York: Prentice-Hall.
- HOLT, ROBERT R. (1956) : Gauging primary and secondary processes in Rorschach responses. *J. proj. Tech.*, 20, 14-25.
- HOLT, ROBERT R. (1958a) : Clinical and statistical prediction: a reformulation and some new data. *J. abnorm. soc. Psychol.*, 56, 1-12.
- HOLT, ROBERT R. (1958b) : Formal aspects of the TAT—a neglected resource. *J. proj. Tech.*, 22, 163-172.
- HOLT, ROBERT R. (1959) : Personality growth in psychiatric residents. *AMA Arch. Neurol. Psychiat.*, in press.
- HOLT, ROBERT R., & LUBORSKY, LESTER B. (1952) : Research in the selection of psychiatrists: a second interim report. *Bull. Menninger Clin.*, 16, 125-135.
- HOLT, ROBERT R., & LUBORSKY, LESTER B. (1955) : The selection of candidates for psychoanalytic training: on the use of interviews and psychological tests. *J. Amer. psychoanal. Ass.*, 3, 666-681.
- HOLTZMAN, W. H., & SELLS, SAUL B. (1954) : Prediction of flying success by clinical analysis of test protocols. *J. abnorm. soc. Psychol.*, 49, 485-490.
- HORN, DANIEL (1943) : An experimental study of the diagnostic process in the clinical investigation of personality. Unpublished doctoral dissertation, Harvard Univer.
- HORST, PAUL, et al. (1941) : *The prediction of personal adjustment*. New York: Social Science Research Council Bulletin, 48.
- HOWE, LOUISA P. (1955) : Some sociological aspects of identification. In W. Muensterberger, & S. Axelrad, Eds., *Psychoanalysis and the social sciences*. Vol. 4. New York: International Universities Press.
- JENKINS, WILLIAM L. (1953) : An index of selective efficiency (S) for evaluating a selection plan. *J. appl. Psychol.*, 37, 78.
- JONES, ERNEST (1946) : A valedictory address. *Int. J. Psychol-Anal.*, 27, 7-12.
- JONES, ERNEST (1948) : *Papers on psychoanalysis*. 5th ed. Baltimore: Williams & Wilkins.
- JONES, ERNEST (1955) : *The life and work of Sigmund Freud*. Vol. 2. "Years of maturity, 1901-1919." New York: Basic Books.
- KASS, WALTER (1952) : Rorschach indications for predicting proficiency in the selection of physicians for training in psychiatry in a residency program. Unpublished doctoral dissertation, New York Univer.
- KELLY, E. L., & FISKE, D. W. (1950) : The prediction of success in the VA training program in clinical psychology. *Amer. Psychologist*, 5, 395-406.
- KELLY, E. L., & FISKE, D. W. (1951) : *The prediction of performance in clinical psychology*. Ann Arbor, Michigan: Univer. of Michigan Press.
- KENDALL, WILLIAM E. (1956) : Industrial psychology. Pp. 197-232 in Farnsworth & McNemar (1956).

- KIHSS, PETER (1945): 10,000 veterans a month return with neuropsychiatric trouble. *New York Herald Tribune*, April 14.
- KLEIN, GEORGE S., & SCHLESINGER, H. J. (1951): Perceptual attitudes toward instability: I. Prediction of apparent movement experiences from Rorschach responses. *J. Pers.*, 19, 289-302.
- KLEIN, HENRIETTE R. (1953): Selection of candidates for training in psychoanalysis. Paper read at the Midwinter Meeting of Amer. Psychoanal. Ass.
- KNIGHT, ROBERT P. (1946): Criteria for judging the suitability of analytic candidates: questionnaire results. Topeka: Menninger Foundation. (Mimeo.)
- [KNIGHT, ROBERT P.] (1947): The selection of candidates for training in psychiatry; philosophy and methods. *Bull. Menninger Clin.*, 11, 83-89.
- KNIGHT, ROBERT P. (1953): The present status of organized psychoanalysis in the United States. *J. Amer. psychoanal. Ass.*, 1, 197-221.
- KUBIE, LAWRENCE S. (1948): in *Training in clinical psychology*. Transactions of the Josiah Macy, Jr. Foundation Conference, March 27-28, New York City. Burlington, Vt.: *J. clin. Psychol., Monogr. Suppl.*, No. 3.
- KUBIE, LAWRENCE S. (1950): *Practical and theoretical aspects of psychoanalysis*. New York: International Universities Press.
- LAMPL-DE GROOT, JEANNE (1954): Problems of psycho-analytic training. *Int. J. Psycho-Anal.*, 35, 184-187.
- LANDSBERGER, HENRY A. (1956): Final report on a research project in mediation. *Labor Law J.*, 7, 501-507.
- LANIER, LYLE H., et al. (1949): Army Air Forces Aviation Psychology Program Research Reports: special reviews. *Psychol. Bull.*, 46, 499-528.
- LEVINE, MAURICE (1952): Principles of psychiatric treatment. Pp. 307-366 in Franz Alexander & Helen Ross, Eds., *Dynamic psychiatry*. Chicago: Univer. of Chicago Press.
- LEWIN, BERTRAM D. (1946): Countertransference in the technique of medical practice. *Psychosom. Med.*, 8, 196-199.
- LEWIS, NOLAN D. C. (1941): *A short history of psychiatric achievement*. New York: Norton.
- LORAND, SANDOR (1955): The teaching of psychotherapeutic technique to residents in psychiatry. *Acta Psychother.*, 3, Suppl., 218-224.
- LUBORSKY, LESTER B. (1945): Aircraft recognition: II. A study of prognostic tests. *J. appl. Psychol.*, 29, 449-457.
- LUBORSKY, LESTER B. (1949): The personality requisites for work in psychiatry as revealed in the TAT. *Amer. Psychologist*, 4, 258 (abstract).
- LUBORSKY, LESTER (1952a): Personality of the psychotherapist. *Menninger Quart.*, 6, (4), 1-6.
- LUBORSKY, LESTER (1952b): The personalities of more and less successful psychotherapists. *Amer. Psychologist*, 7, 337 (abstract).
- LUBORSKY, LESTER (1953): Self-interpretation of the TAT as a clinical technique. *J. proj. Tech.*, 17, 217-223.
- LUBORSKY, LESTER (1954): Selecting psychiatric residents: survey of the Topeka research. *Bull. Menninger Clin.*, 18, 252-259.

## 372 Bibliographic References

- LUBORSKY, LESTER, & HOLT, ROBERT R. (1957) : The selection of candidates for psychoanalytic training. Implications from research on the selection of psychiatric residents. *J. clin. exp. Psychopath. Quart. Rev. Psychiat. Neurol.*, **18**, 166-176.
- LUBORSKY, LESTER, HOLT, ROBERT R., & MORROW, WILLIAM R. (1950) : Interim report of the research project on the selection of medical men for psychiatric training. *Bull. Menninger Clin.*, **14**, 92-101.
- LUNDBERG, G. A. (1941) : Case studies vs. statistical methods—an issue based on misunderstanding. *Sociometry*, **4**, 379-383.
- MACKINNON, DONALD W. (1951) : The effects of increased observation upon the accuracy of prediction. *Amer. Psychologist*, **6**, 311 (abstract).
- MALAMUD, WILLIAM (1944) : The history of psychiatric therapies. Pp. 273-324 in Hall, J. K. (1944).
- MASLOW, ABRAHAM (1954) : *Motivation and personality*. New York: Harper.
- MAY, ROLLO (1950) : The work and training of the psychological therapist. *Psychol. Serv. Cent. J.*, **2**, 3-23.
- MAYMAN, MARTIN (1953) : A study of the sense of self. Unpublished doctoral dissertation, Kansas Univer.
- MCAULIFFE, CHARLES (1956) : Clinical versus actuarial prediction. Pp. 99-106 in *Proceedings, 1955 invitational conference on testing problems*. Princeton, N.J.: Educational Testing Service.
- MEEHL, PAUL E. (1954) : *Clinical vs. statistical prediction*. Minneapolis: Univer. of Minnesota Press.
- MEEHL, PAUL E. (1956a) : Clinical vs. actuarial prediction. Pp. 136-141 in *Proceedings, 1955 invitational conference on testing problems*. Princeton, N.J.: Educational Testing Service.
- MEEHL, PAUL E. (1956b) : Wanted—a good cookbook. *Amer. Psychologist*, **11**, 263-272.
- MEEHL, PAUL E., & ROSEN, ALBERT (1955) : Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychol. Bull.*, **52**, 194-216.
- MELTON, RICHARD S. (1952) : A comparison of clinical and actuarial methods of prediction with an assessment of the relative accuracy of different clinicians. Unpublished doctoral dissertation, Univer. of Minnesota.
- MENNINGER, KARL A., et al. (1945) : Psychiatric education issue. *Bull. Menninger Clin.*, **9**, 29-70.
- MENNINGER, KARL A., et al. (1947) : The Menninger Foundation School of Psychiatry. (Special issue.) *Bull. Menninger Clin.*, **11**, 77-108.
- MENNINGER, KARL A. (1950) : A psychiatric fable. *Bull. Menninger Clin.*, **14**, 129-130.
- MENNINGER, KARL A. (1952) : What are the goals of psychiatric education? *Bull. Menninger Clin.*, **16**, 153-158.
- MENNINGER, KARL A. (1957) : Psychological factors in the choice of medicine as a profession. (Parts I and II.) *Bull. Menninger Clin.*, **21**, 51-58, 99-106.
- MENNINGER, WILLIAM C. (1948a) : *Psychiatry in a troubled world*. New York: Macmillan.

- MENNINGER, WILLIAM C. (1948b) : *Psychiatry: its evolution and present status*. Ithaca, N.Y.: Cornell Univer. Press.
- MERTON, R. K., BLOOM, S., & ROGOFF, N. (1956) : Studies in the sociology of medical education. *J. med. Educ.*, 31, 552-565.
- MILLS, C. WRIGHT (1953) : *White collar*. New York: Oxford Univer. Press.
- MODLIN, HERBERT C. (1955) : An evaluation of the learning process in a psychiatric residency program. *Bull. Menninger Clin.*, 19, 135-159.
- MORENO, J. L. (1953) : *Who shall survive? Foundations of sociometry, group psychotherapy and sociodrama*. 2nd ed. Beacon, N.Y.: Beacon House.
- MORRIS, W. W. (1951) : Validity of the Professional Aptitude Test in medicine. *J. med. Educ.*, 26, 56-58.
- MUNCIE, WENDELL, & BILLINGS, EDWARD G. (1951) : A survey of conditions of private psychiatric practice throughout the United States and Canada. *Amer. J. Psychiat.*, 108, 171-172.
- MÜNSTERBERG, HUGO (1913) : *Psychology and industrial efficiency*. Boston: Houghton Mifflin.
- MURPHY, GARDNER (1947) : *Personality. A biosocial approach to origins and structure*. New York: Harper.
- MURPHY, LOIS B., & LADD, HENRY (1953) : *Emotional factors in learning*. New York: Columbia Univer. Press.
- MURRAY, HENRY A., et al. (1938) : *Explorations in personality*. New York: Oxford Univer. Press.
- National Committee for Mental Hygiene (1945) : Mental hygiene committee comments on psychiatric facilities for discharged veterans. *J. Amer. med. Ass.*, 128, 744.
- National Mental Health Committee (1956) : *What are the facts about mental illness in the United States?* Washington, D.C.: the Committee.
- NUNBERG, HERMAN (1948) : *Practice and theory of psychoanalysis*. New York: Nervous and Mental Disease Monographs.
- OBERNDORF, CLARENCE P. (1953) : *A history of psychoanalysis in America*. New York: Grune & Stratton.
- Office of Strategic Services, Assessment Staff (1948) : *Assessment of men; selection of personnel for the Office of Strategic Services*. New York: Rinehart.
- PACE, C. ROBERT (1939) : Factors influencing questionnaire returns from former university students. *J. appl. Psychol.*, 23, 388-397.
- PATERSON, DONALD G., & DARLEY, JOHN G. (1936) : *Men, women and jobs*. Minneapolis: Univer. of Minnesota Press.
- PFISTER, OSKAR (1917) : *The psychoanalytic method*. London: Kegan Paul.
- PIAGET, JEAN (1952) : *The origins of intelligence in children*. New York: International Universities Press, Inc.
- PORTER, W. C., & DAVIDSON, H. A. (1947) : Alumni appraisal of psychiatric education. *Amer. J. Psychiat.*, 103, 440-445.
- Psychiatry and Neurology Division, Department of Medicine and Surgery, Veterans Administration (1953) : *Planning in psychiatry, neurology, and*

- clinical psychology. A forecast: 1953-1960. Washington, D.C.: Veterans Administration.
- RAPAPORT, DAVID (1942): Principles underlying projective techniques. *Charact. & Pers.*, 10, 213-219.
- RAPAPORT, DAVID (1946): Principles underlying non-projective tests of personality. *Ann. N.Y. Acad. Sci.*, 66, 643-652.
- RAPAPORT, DAVID (1950a): The theoretical implications of diagnostic testing procedures. Pp. 241-271 in *Rapports, Congrès international de psychiatrie*. Vol. II. Paris: Hermann.
- RAPAPORT, DAVID (1950b): On the psychoanalytic theory of thinking. *Int. J. Psycho-Anal.*, 31, 1-10.
- RAPAPORT, DAVID (1951): *Organization and pathology of thought*. New York: Columbia Univer. Press.
- RAPAPORT, DAVID, GILL, MERTON M., & SCHAFER, ROY (1945-46): *Diagnostic psychological testing*. Chicago: Yearbook Publishing Co. 2 vols.
- REDLICH, FREDERICK (1950): The psychiatrist in caricature: an analysis of unconscious attitudes toward psychiatry. *Amer. J. Orthopsychiat.*, 20, 560-571.
- REDLICH, FRITZ (1953): See Bingham (1953).
- REIK, THEODOR (1948): *Listening with the third ear*. New York: Farrar, Straus.
- Research Staff of the National Manpower Council (1953): *A policy for scientific and professional manpower*. New York: Columbia Univer. Press. Summarized as: National Manpower Council. *Science*, 117, 617-622.
- ROGERS, CARL R. (1942): *Counseling and psychotherapy*. New York: Houghton Mifflin.
- ROHLES, FREDERICK H., JR. (1953): Construction and validation of a test battery as determinant of overseas adjustment potential. Validation of a biographical inventory of select airmen who desire to extend their Alaskan duty tour. Arctic Aeromedical Laboratory Project No. 22-0201-0002, research report No. 1, Series 2. Ladd Air Force Base, Alaska.
- RUESCH, JURGEN (1953): Social factors in therapy. Chap. VI in *Psychiatric treatment*. *Proc. Assoc. Res. nerv. ment. Dis.*, Vol. 31. Baltimore, Md.: Williams & Wilkins.
- RUESCH, JURGEN, & BATESON, GREGORY (1951): *Communication: the social matrix of psychiatry*. New York: Norton.
- SACHS, HANNS (1947): Observations of a training analyst. *Psychoanal. Quart.*, 16, 157-168.
- SARBIN, THEODORE (1941): The relative accuracy of clinical and statistical predictions of academic success. Unpublished doctoral dissertation, Ohio State Univer.
- SARBIN, THEODORE R. (1942): A contribution to the study of actuarial and statistical methods of prediction. *Amer. J. Social.*, 48, 593-603.
- SARBIN, THEODORE R. (1944): The logic of prediction in psychology. *Psychol. Rev.*, 51, 210-228.
- SCHAFER, ROY (1948): *The clinical application of psychological tests*. New York: International Universities Press.

- SCHAFER, ROY (1949) : Psychological tests in clinical research. *J. consult. Psychol.*, 13, 328-334. Reprinted in R. P. Knight, & C. R. Friedman, Eds., *Psychoanalytic psychiatry and psychology: clinical and theoretical papers, Austen Riggs Center*. Vol. I. New York: International Universities Press, 1954.
- SCHAFER, ROY (1954) : *Psychoanalytic interpretation in Rorschach testing*. New York: Grune & Stratton.
- SCHWARTZ, EMANUEL K., & ABEL, THEODORA M. (1955) : The professional education of the psychoanalytic psychotherapist. *Amer. J. Psychother.*, 9, 253-261.
- SEVENTH MEDICAL ECONOMICS SURVEY (1952-53) : Your economic weather vane. *Medical Economics*, 20, (2), 85-96; (8), 117-124.
- SHAKOW, DAVID, et al. (1947) : Recommended graduate training program in clinical psychology. *Amer. Psychologist*, 2, 539-558.
- SHARPE, ELLA (1950) : *Collected papers on psychoanalysis*. London: Hogarth Press.
- SHEEHAN, JOSEPH G. (1953) : Rorschach changes during psychotherapy in relation to personality of the therapist. *Amer. Psychologist*, 8, 434 (abstract).
- STALNAKER, JOHN M. (1950) : Medical College Admission Test. *J. Ass. Amer. Med. Coll.*, 25, 428-434.
- STALNAKER, JOHN M. (1954) : Personal communication.
- STEINZOR, BERNARD (1953) : Patterns of academic achievement and personality. Unpublished manuscript.
- STRONG, EDWARD K. (1943) : *Vocational interests of men and women*. Stanford, California: Stanford Univer. Press.
- STRONG, EDWARD K. JR. & TUCKER, ANTHONY C. (1952) : The use of vocational interest scales in planning a medical career. *Psychol. Monogr.*, 66, No. 9 (Whole No. 341). Pp. vi, 61.
- STUIT, DEWEY B., & SCHLICHTER, R. J. (1948) : *Handbook for advisers to students planning to enter medicine*. Chicago, Ill.: Association of American Medical Colleges.
- STUIT, DEWEY B., et al., (1949) : *Predicting success in professional schools*. Washington: American Council on Education.
- SZONDI, LIPOT (1952) : *Experimental diagnostics of drives*. New York: Grune & Stratton.
- SZUREK, S. A. (1949) : Remarks on training for psychotherapy. *Amer. J. Orthopsychiat.*, 19, 36-51.
- TAFT, RONALD (1955) : The ability to judge people. *Psychol. Bull.*, 52, 1-23.
- TAYLOR, FREDERICK W. (1913) : *Principles of scientific management*. New York: Harper.
- TERMAN, LEWIS M. (1954) : The discovery and encouragement of exceptional talent. *Amer. Psychologist*, 9, 221-230.
- THOMPSON, CLARE W., & MACDONALD, MARTHA W. (1953) : Personality characteristics of residents in psychiatry. *Amer. Psychologist*, 8, 446-447.
- THORNDIKE, ROBERT L. (1933) : The effect of the interval between test and re-test on the constancy of the IQ. *J. educ. Psychol.*, 24, 543-549.

## 376 Bibliographic References

- THORNDIKE, ROBERT L. (1949): *Personnel selection, test and measurement techniques*. New York: Wiley.
- TOMKINS, SILVAN S. (1947) : *The Thematic Apperception Test*. New York: Grune & Stratton.
- TRYTTEN, M. H. (1957) : Soviet and U.S. professional and technical manpower. *Science*, 126, 11-15.
- TURNER, E. L., WIGGINS, W. S., & TYNIER, ANNE (1955) : Medical education in the United States and Canada. *J. Amer. med. Ass.*, 159, 563-606.
- UNITED STATES CONGRESS (1953) : *Health inquiry into the causes, control and remedies of the diseases of mankind: part 4*. Hearings before the House Committee on Interstate and Foreign Commerce, 83rd Congress. Washington, D.C.: U. S. Government Printing Office.
- VITELES, MORRIS S. (1934) : *The science of work*. New York: Norton.
- WALLERSTEIN, ROBERT S., & ROBBINS, LEWIS L. (1956) : The psychotherapy research project of the Menninger Foundation. IV. Concepts. *Bull. Menninger Clin.*, 20, 239-262.
- WARD, CLYDE H. (1954) : Psychiatric training in university centers. *Amer. J. Psychiat.*, 111, 123-131.
- WECHSLER, DAVID (1939) : *The measurement of adult intelligence*. Baltimore, Md.: Williams & Wilkins.
- WEIGERT, EDITH (1952) : Contribution to the problem of terminating psychoanalysis. *Psychoanal. Quart.*, 21, 465-480.
- WEIR, JOHN R. et al. (1956) : Report on technical manpower requirements, supplies and limits. In California Institute of Technology, *Speculative projection of civilization in the twenty-first century*. Summarized in *New York Times*, May 21, p. 19.
- WELD, H. P. & ROFF, M. (1938) : A study in the formation of opinion based upon legal evidence. *Amer. J. Psychol.*, 51, 609-628.
- WHITE, ROBERT W. (1952) : *Lives in progress*. New York: Dryden.
- WHITE, WILLIAM A. (1925-26) : Presidential address. *Amer. J. Psychiat.*, 5, 1-20.
- WHITEHORN, JOHN C., et al., Eds. (1952) : *Psychiatry and medical education*. Report of the 1952 Conference on Psychiatric Education held at Cornell Univer. Washington, D.C.: American Psychiatric Association.
- WHITEHORN, JOHN C., et al., Eds. (1953) : *The psychiatrist, his training and development*. Report of the 1952 Conference on Psychiatric Education held at Cornell Univer. Washington, D.C.: American Psychiatric Association.
- WILLIAMS, STANLEY B., & LEAVITT, HAROLD J. (1947) : Group opinion as a predictor of military leadership. *J. consult. Psychol.*, 11, 283-291.
- WINSLOW, WALKER (1956) : *The Menninger story*. New York: Doubleday.
- WITTMAN, M. PHYLLIS (1941) : A scale for measuring prognosis in schizophrenic patients. *Elgin Papers*, 4, 20-33.
- ZILBOORG, GREGORY, & HENRY, GEORGE W. (1941) : *A history of medical psychology*. New York: Norton.

## INDEX

- Abbott, Dr., 55, 159-97, 198, 244, 250, 287  
Abel, Theodora, 214 *n.*  
abreaction index, 280  
ACE intelligence test, 104  
actuarial system, 79-89, 100, 231-33  
administrative competence, 154, 169 f.,  
    174, 176, 201  
    definition of, 140  
Admissions Committee, 34, 57, 118, 325  
    decisions of, 130-31, 132-34  
    members of, 329  
    pressures on, 297-98  
    procedures of, 129, 289, 327-29  
Adorno, T. W., 226, 275  
age as related to competence, 223, 263,  
    270, 330  
Air Force studies, 105  
    (see also Army Air Force)  
Albee, George W., 353  
    *quoted*, 354  
Alexander, Franz, 357  
Allport, Gordon W., 92, 239  
American Association of Medical Col-  
    leges, 84  
American Board of Psychiatry and Neu-  
    rology, 20, 109, 116 f., 225  
    certification by, 119-20, 131  
    requirements of, 15  
American Medical Association, 16, 355  
American Psychiatric Association, 9, 256  
American Psychoanalytic Association,  
    338, 340  
American Psychological Association, 265  
Anderson, Donald G., 13  
Anderson, R. C., 22  
anxiety, 263, 279, 282  
appearance, physical, 264, 271, 291  
    as source of predictive error, 235-38,  
        291  
Army Air Force tests, 87-88, 95 f. (see also  
    Air Force)  
Army Alpha test, 80 f.  
Army General Classification Test, 87  
assessment procedures, 140, 151-54  
authoritarianism, 247 f., 262, 263, 280,  
    289 (see also E and F scales, ethno-  
    centric prejudice)  
Barron, Frank, 335  
"basic trust," 313 *n.*  
Bellevue Scale of Adult Intelligence (see  
    Wechsler-Bellevue Scale)  
Bellinger, Clarence, *quoted*, 257  
Berengarten, Sidney, 7  
Bernreuter Personality Inventory, 83  
Bettelheim, Bruno, 153 *n.*  
Bingham, Walter V., *quoted*, 218  
Binet, Alfred, 80  
Binet tests, 80  
Blain, Daniel, 19  
Blenker, M., 106  
Blitzsten, Lionel, *quoted*, 300-301  
Boothe, Bert E., xii  
Boston Psychopathic Hospital, 5 *n.*  
Bowman, Karl M., *quoted*, 260  
Boys' Industrial School, Kansas, 25 .  
Bradley, General Omar N., 10  
Brooklyn State Hospital, 5 *n.*

- Brown, Kathryn, xi  
 Bureau of Salesmanship Research, 81  
 Burgess, E. W., 86
- California Ethnocentrism Scale, (*see* E scale)  
 California Institute of Technology, 355  
 California Personality Inventory, 335  
 Carnegie Institute of Technology, 81  
 Casey, Jesse F., xii, 22  
 Cattell, James McKeen, 79  
 certification, 71, 116, 119-20, 131  
 Chappell, Philip, xi  
 Chess, Rudolph, 20  
 chi-square, 363  
 child psychiatry (*see* Southard School)  
 Cincinnati (*see* University of Cincinnati)  
 City-County Clinics, 25  
 clarity of thought, 275, 277, 287, 322  
     defined, 145  
 clinical psychology, 7 f., 41, 97, 224-25,  
     265  
     training programs in, 11, 16, 20  
 clinical tradition of selection, 89-99  
 College Entrance Examination Board, 86 (*see also* Educational Testing Service)  
 Colorado (*see* University of Colorado)  
 Columbia Psychoanalytic Institute, 338 ff.  
 Committee on Medical Education, 256,  
     351  
 competence (*see* administrative; diagnostic; management; over-all; psychotherapeutic)  
 Conference on Graduate Psychiatric Education, 256  
 Conference on Post-war Problems of Psychoanalytic Training, 338  
 contamination of predictions, 158, 178,  
     245  
 correlation, explanation of, 359-61  
     multiple, 87 f., 231 f., 363  
 Cottrell, Leonard S., Jr., 86  
 Coyne, LolaFaye, xi  
 credentials, 141, 154, 155, 158, 214, 309,  
     312, 328 f., 334  
     manual, validity of, 151  
 criteria  
     actually used, 116-27  
     definition of, 108  
     dichotomous vs. scaled, 112-14  
     factor analysis of, 126  
     follow-up, 134  
     intensive work recommended, 333-34  
     in-training, 110-12, 133  
     Michigan Project, 97  
     post-training, 110-12, 132  
     problem of, 108-27  
     qualitative, 126-27  
     ratings vs. special measures, 114-16  
     relationships among the various, 127  
     recommended system for collecting data on, 345  
     reliabilities of the, 125-26  
     stable, need for, 336  
     types of, 109-16  
     (*see also* administrative competence; diagnostic competence; management competence; over-all competence; psychotherapeutic competence)  
 criterion groups, personality differences between, 276-87  
 criterion information, specimen analysis of (Dr. Abbott), 184-95  
 Cronbach, Lee J., xii, 101, 104 f., 209, 348  
 cross-validation, 148, 150-51, 198-99, 334,  
     348 (*see also* manuals)  
 cues, 147, 199, 203, 334  
 cue-sums, 150 f., 199-200, 201, 203, 205,  
     208 f.  
 cultural wealth, 241, 257, 262, 266, 270,  
     275, 288, 290 f., 323  
     defined, 145  
 curiosity, psychological, 264, 265, 290-91,  
     308  
     defined, 143  
 cutting points and criteria, 112 ff.  
 cutting score, 200-203, 206, 213, 229, 232
- Darley, John G., 82  
 defenses, value of, 299  
 Design I (rule-of-thumb), 208, 210, 212-  
     15, 249-51, 348  
 Design II (method of guided clinical judgment), 159, 198, 203, 206, 210,  
     212-15, 348  
 limitations of, 249-51  
 plans for, 137-50  
 results of, 198-215  
     (*see also* predictive study)
- Deutsch, Albert, 7 n., 9, 11

- development, ability for, 281-87  
 diagnostic competence, 154, 169 f., 174,  
   176, 201 ff., 206, 210, 225, 228  
   definition of, 139  
 Dickey, Marguerite, 353  
   *quoted*, 354  
 "doctor draft," 16  
 "Dr. Karl's Colloquium," 30  
 drop-outs, 60 f., 118 f., 127, 149  
 Du Bois, Phillip H., 88 n.
- E Scale, 226, 317  
 education, 85-86, 89  
   psychiatric, 13-17  
 Educational Testing Service, 84, 86  
 "ego-strength" scale, 335  
 Ehrenrich, Gerald, 339  
 Ekstein, Rudolf, xi f., 340  
 Elgin State Hospital, 5 n.  
 emotional appropriateness, 161, 166, 177,  
   274 f., 279, 292, 320-22  
   defined, 144  
 emotional control, adequate, 55, 161, 166,  
   248, 258, 271, 279 f., 320 f.  
   defined, 144  
 emotions, handling of, 144, 320-22  
 empathy, 155, 166, 237 f., 256, 267, 274,  
   277, 287, 315-16  
   defined, 144  
 endowment, 278  
 Erikson, Erik H., 98, 313 n., 320  
 Eron, Leonard D., 66  
 errors (*see* predictive errors)  
 Escalona, Sibylle K., x  
 ethical standards, well vs. poorly inter-  
   nalized, 266, 271, 275, 282, 287 f.,  
   323 f.  
   defined, 145  
 ethnocentric prejudice, freedom from,  
   269, 275, 288, 316-17  
   defined, 144  
   (*see also* E Scale)  
 evaluations (*see* peers' evaluations; super-  
   visors' evaluations)  
 Ewalt, Jack, *quoted*, 257  
 examinations, 62-63, 115, 133  
 extraversion (*see* introversion-extraver-  
   sion)
- F Scale, 226  
 façade (*see* genuineness vs. façade)
- "face validity," 127, 298  
 factor analysis, 126, 211, 251, 364  
 fascist attitudes (*see* F Scale; authoritar-  
   ianism)  
 federal aid to medical education, 355  
 Fellows (*see* Menninger Fellows)  
 Fiedler, Fred, 280  
 Fine, Benjamin, 354  
 Finesinger, J. E., 102  
 Fiske, Donald W., 7, 41, 65, 96 f., 116,  
   139, 219, 266, 279 f., 348  
 Flanagan, John C., 140  
 follow-up, 111, 132, 134, 336  
 foreign applicants, 331-32  
 formal aspects of TAT (*see* TAT man-  
   uals)  
 Frank, Jan, xi  
 Freud, Sigmund, 49, 90, 98, 110, 290, 306,  
   337
- Galton, Francis, 79 f.  
 genuineness vs. façade, 163, 166, 273, 325  
   defined, 145  
 Gestalt psychology, 92  
 Giedt, F. Harold, 238  
 Gilbreths, F. and L., 80  
 Gildea, Edwin F., 40  
   *quoted*, 257  
 Gildea, Margaret C. L., 40  
   *quoted*, 40  
 Gill, Merton M., xi  
 Gough, H. G., 335  
 Gough, Twila, xi  
 Graduate Record Examination, 84  
 Greenhill, Maurice H., 256, 351  
 Gregg, Alan, *quoted*, 54, 69  
 Grinker, Roy R., *quoted*, 257  
 Group for Advancement of Psychiatry,  
   256, 351  
 group therapy, 61  
 growth (*see* personality growth, capacity  
   for)
- Hacker, Frederick, xi  
 Hall, Bernard, 8, 14, 33  
 Hall, J. K., 7 n.  
 "halo effect," 216, 218  
 Hanfmann, Eugenia, 96  
 Hartmann, Heinz, 98  
 Harvard Psychological Clinic, 92 ff.

- Hawley, General Paul R., 10  
 Hedstrom, George, xii  
 Heron, M. A., 85  
 historical background, 7-17  
 Hollingsworth, Irene, xi  
 Holt, Robert R., 52, 102, 142, 146, 152,  
     326, 340  
 Holzman, Philip, xi  
 honesty, 145, 323-25  
 Horn, Daniel, 92 n., 156 n., 239  
 Horst, Paul, 85 f., 99  
 hospitals (*see* state hospitals; VA hospitals)  
 hostility, adequate handling of, 161, 166,  
     171, 274 f., 282 f., 291 f., 320 f.  
     defined, 144  
 humor, sense of, 177, 262, 265, 268, 276,  
     290, 311-12  
     defined, 143  
     \*  
 identity, sense of, 320  
 impulse and affect (*see* emotions, handling of; emotional control)  
 income, 43, 75-76  
 "incompetents," 58-59, 76  
 industrial psychology, 80 ff.  
 inhibition, 235, 241, 244, 248, 282, 293  
 insight (*see* self-objectivity)  
 Institute of the Pennsylvania Hospital,  
     5 n.  
 integrity (*see* ethical standards)  
 intellectual effectiveness, 145, 322-33 (*see also* IQ, Verbal IQ)  
 interest in psychiatry (and in work), 59,  
     155, 176, 266, 269, 274 f., 277 f.,  
     288, 309 f. (*see also* Strong Vocational Interest Blank, Psychiatrist A key)  
     defined, 143  
     origins of, 44-51  
 interests, vocational (*see* Strong Vocational Interest Blank)  
 interview, 5, 61, 83, 128-29, 143, 151-52,  
     155, 309, 312, 320, 347 f.  
 Blitzsten's procedure, 300-301  
 clinical, 90  
 complaints about, 319  
 evaluative criterion, 139  
 manual, 147, 150, 152, 200, 203-204,  
     209  
 recommendations for, 298-301  
 specimen analysis of (Dr. Abbott),  
     175-76  
     "stress," 299  
 structured, 141  
 terminal, 61  
 transcribed, 47-50  
 validity of predictions from, 130-35,  
     203 f., 206, 209, 212-15  
 interviewers  
     aptitude of, 329  
     comments of, 241  
     errors of, 234-38, 250  
     improving the skill of, 239-40  
     number of, 135, 329  
 introversion-extroversion, 39 f., 241, 247,  
     257, 264, 271  
 intuition, 256, 263, 267, 288, 289  
 IQ of residents, 40 (*see also* Verbal IQ)  
 "Irresponsibles," 59 f.  
 James, William, 94  
 job analysis, 137, 139-40, 333  
 "journeymen," 56-58, 60  
 Jung, Carl, 92  
 Kansas (*see* University of Kansas)  
 Kansas City mental hygiene clinic, 25  
 Kass, Walter, xi, 146, 202  
 Keller, Helen, 30  
 Kelly, E. L., 7, 41, 65, 96 f., 116, 139, 219,  
     266, 279 f., 348  
 Klein, George S., xii  
 Klein, Henriette R., 7, 338, 340  
 Knight, Robert P., x ff., 4, 257, 261, 300  
 Knox, James, xi  
 Kubie, Lawrence S., 44 n., 353 n.  
 Kuder Preference Test, 83  
 Landsberger, Henry A., 202, 352  
 Langley Porter Clinic, 5 n.  
 "late developers," 54-56, 60, 68 n., 76, 185  
 letters of recommendation, 5, 83-84, 154,  
     319  
 Levinson, Daniel, 226  
 Life Insurance Sales Research Bureau, 81  
 liking, as a predictor, 169 f., 216-21, 279  
 Luborsky, Lester, 142, 146, 153 n., 335,  
     340  
 Lundberg, G. A., 100

- McArthur, Charles, 106  
 McDougall, William, 92  
 MacKinnon, Donald, 96  
 Maine, Harold, 11  
 management of patients, competence in, 154, 169 f., 174, 176, 211  
 definition of, 139  
 manuals to guide analysis of tests and interview, 334 (*see also* Rorschach, TAT, Self-Interpretation, Interview, PRT, Credentials)  
 development of, 142, 146-47  
 function of, 199, 208  
 method (*see* Design II)  
 reliability of, 199-200  
 scoring of, 156-57  
     applied to Dr. Abbott, 165, 171-72,  
     175  
 validities of, 150-51, 200-204  
     compared to free ratings, 208-209  
 marital adjustment studies, 86, 223-24  
 material preoccupations, 143, 275, 288, 304 f. (*see also* monetary motivation)  
 Marquis, Donald, xi  
 MAT (*see* Miller Analogies Test)  
 Mayman, Martin, xi, 278  
 Medical Aptitude Test, 84  
 Medical College Admission Test, 85  
 medical school record, 222-23, 257, 330  
 Meehl, Paul E., 91 n., 100 f., 103 ff., 209  
 Melton, Richard S., 104  
 Menninger, Charles F., 18  
 Menninger, Karl A., x ff., 4, 18 f., 30, 49,  
     57, 255 ff., 260  
     *quoted*, 258, 259-60  
 Menninger, William C., x f., 7 n., 20, 87  
     *quoted*, 65  
 Menninger Clinic, 18, 23, 25, 125, 185  
 Menninger Fellows, 23, 34 (*see also* residents)  
 Menninger Foundation, x f., 4, 5 n., 18 ff., 23, 26 f., 64, 129  
 Menninger Hospital (Sanitarium), 18, 25  
 Menninger Monographs Committee, x  
 Menninger School of Psychiatry, ix f.  
     alumni (*see* residents)  
     application form, 309  
     clinical assignments, 24-26  
     completion of residency at, 117 f.  
     conceptions of, held by applicants, 33-34  
     curriculum, 23-24  
     formal examinations at, 62  
     history of, 4, 11, 18-32  
     and medical school record, 222-23  
     minority-group representation at, 39  
     objectives of, 109  
     problem of selection at, 113  
     psychoanalytic orientation of, 33 f., 351  
     representativeness of, 351  
     representatives visiting medical schools, 14, 354  
     size of, 352  
     tests used at, 5  
     trainees 1946-51, 276  
     (*see also* Admissions Committee)  
     mental health of residents, 60 f., 250, 257,  
     271, 280-81  
     mental illness, definition of, 11  
     Meyer, Adolf, 63  
     Michael Reese Hospital-Associated Facilities, 5 n.  
     Michigan (*see* University of Michigan)  
     Michigan Project, 41, 92 n., 96-97, 139,  
     158 n., 219, 279, 348 f.  
     “middle-of-the-roaders,” 56  
     Mill, John Stuart, *quoted*, 266  
     Miller, James G., xi, 96  
     Miller Analogies Test (MAT), 5, 230  
     Minnesota (*see* University of Minnesota)  
     Minnesota Employment Stabilization Research Institute, 82  
     Minnesota Multiphasic Personality Inventory, 5, 146, 335  
     minority groups, 39  
     Mitchell, S. Wier, 40  
     Modlin, Herbert C., xi, 62  
     monetary motivation, 43, 304-305 (*see also* material preoccupations)  
     Morris, W. W., 84  
     Morrow, William R., x, 142, 146, 226  
     motivations for psychiatric work, 43-51,  
     143, 290, 304-310  
     multiple regression (*see* correlation, multiple)  
     Münsterberg, Hugo, 80  
     Murphy, Gardner, x, xii  
     Murray, Henry A., 92 ff., 100, 152, 256,  
     258  
     *quoted*, 260

- musicians, interest patterns of, 41
- National Institute of Mental Health, 16 f.
- National Manpower Council, 355
- National Mental Health Act (1946), 16
- need to help ("nurturance"), 162 f., 165, 176, 269, 274, 290 f., 306  
defined, 143
- "neurotically hampered," 55
- New York Foundation, x
- New York Hospital, 5 n.
- New York State Psychiatric Institute, 5 n.
- Norristown State Hospital, 5 n.
- North Little Rock V. A. Hospital, 5 n.
- Nunberg, Herman, 337
- nurses, relationships with, 271, 319 (*see also* social relationships with co-workers)
- objectivity toward authority figures, 166, 168, 177, 275, 288, 313 f.  
defined, 144
- objectivity toward dependent figures, 275, 288, 313-15  
defined, 144
- observer agreement, 199
- obsessiveness (*see also* overobsessiveness)
- Occupational Research Program, 82
- Office of Strategic Services (OSS), 82, 92 n., 94 ff.
- Oppenheim, Sadi, 214 n.
- orientation, self-, 143-44, 310-13  
toward others, 144, 313-17
- Osler, Sir William, 49
- outpatient clinic, 191
- over-all competence, 124 ff., 138, 154, 169 f., 174, 176, 203, 207 ff., 212, 214, 219, 222 f., 228, 231 f., 250, 273, 275 f., 288, 292, 317
- Dr. Abbott and, 185  
definition of, 140
- "overestimated," 54 f., 76
- overobsessiveness, 161, 248  
defined, 145
- Parsons, Talcott, xi
- Paterson, Donald G., 82
- pathological tendencies, 65 f., 145, 271, 325-26 (*see also* mental health)
- patients, relationships with, 207-208, 210  
treatment of, as persons, 14
- (*see also* management of patients; objectivity toward dependent figures)
- peers' evaluations, 62, 120 ff., 126 f., 133, 185, 200 ff., 206, 209, 214, 217, 219, 225, 228, 279
- personal history, 37-40, 222-24, 281
- personality differences between better and poorer residents, 272-93
- personality growth, capacity for, 166, 260 f., 268, 275, 281 f., 288, 312-13  
defined, 144
- personality inventories, 82 f., 93, 335 (*see also* Minnesota Multiphasic Personality Inventory, Strong Vocational Interest Blank, tests)
- personality ratings, 176-77
- personality requisites for psychiatry, 255-71  
differences between psychiatrists, psychotherapists, psychoanalysts, 261-65
- specificity to psychiatry, 265-67
- personnel psychology, 80 f., 87
- Picture Reaction Test (PRT), 141, 142-43, 153-54, 155, 214  
manual, 147, 150, 200, 203 f., 208
- Picture-Story Matching Test, 150 n.
- Plato, quoted, 79
- poise, as source of error, 235-38
- practice, psychiatric  
expectations regarding, 42-44  
kinds of, 73-75  
private, 14, 42 f., 72 f., 110 f.  
settings of, 72-73
- predictions  
actuarial, 231-33  
by free ratings, 204-208  
level of generality of, 210-12  
reliability of, 146  
rule-of-thumb, 205 (*see also* Design I)  
specimen analysis of, 159-97  
criterion information, 184-95  
postscript (Dr. Abbott), 195-97  
predictive process, 160-84  
validity of, 130-36, 208-12
- predictive errors, 234-51  
improving the skill of interviewers, 239-40  
in interviewing approach, 234-38  
judgeability of the applicant, 238-39

- retrospective studies, 240-49  
 predictive process, as applied to Dr. Abbott, 160-84  
 predictive study (*see also* Design II)  
     assessment procedures in, 151-54  
     experimental design of, 148-49  
     limitations of, 249-51  
     objectives of, 198-99  
 predictors, validities of miscellaneous facts of personal history, 222-24  
     liking, 216-21  
     personality variables, 273-76  
     Rorschach scores, 230-31  
     self-administrative inventories, 224-27  
 Wechsler-Bellevue measures, 227-30  
 process notes, 284 f.  
 projection, 263, 291, 326  
     defined, 145  
 PRT (*see* Picture Reaction Test)  
 Pruyser, Paul W., xii  
 psychiatric aides, 8, 20  
 psychiatric education, 13-17  
 psychiatric nursing, 9, 16  
 psychiatric social work, 7 ff., 16, 20  
 psychiatrists, personality differences from psychotherapists and psychoanalysts, 261-63, 264-65, 287-93  
 psychoanalysis, 23, 42, 72, 125, 337-46  
 psychoanalysts, personality differences from psychiatrists and psychotherapists, 263-65, 287-93  
 psychological-mindedness, 162, 165, 241, 247, 257, 267, 274 f., 288 f., 307 f.  
     defined, 143  
 psychotherapeutic competence, 124-25, 154, 169 f., 174, 175-76, 201-203, 206, 210 f., 213, 214, 225, 276 n., 317  
     of Dr. Abbott, 185, 188-92  
     definition of, 139  
 psychotherapists, personality characteristics of, 276  
     differences from psychiatrists and psychoanalysts, 263, 264-65, 287-93  
 psychotherapy, 23, 42, 61, 74, 159, 284-86  
     Dr. Abbott, 189  
     caliber of applicants for, 67-68  
     results of, 68-69  
     undergone by residents, 55, 63-69, 358  
 questionnaires  
     attitude, psychiatric, 226-27  
     follow-up, 180  
     self-concept, 278  
 Supplementary Face Sheet, 37 n., 150, 214, 278  
     (*see also* tests)  
 racial prejudice (*see* ethnocentric prejudice; E Scale)  
 Ramana, C. V., xi, 142  
 Rapaport, David, x, xii, 4, 30, 97, 121, 142, 151, 321  
     ratings, predictive, 128-29, 154-55  
     of Dr. Abbott, 168-70, 173-76  
     (*see also* predictions; validities)  
 recommendations, 297-346  
     general, 298-326  
     on interviewing, 298-301  
     psychoanalytic selection, 343-46  
     qualities to look for, 303-26  
     research, 332-36  
     specific, to selectors, 326-32  
     use of the test battery, 302-303  
 recruitment, 113, 353, 355  
 Redlich, Fritz, 256  
     *quoted*, 257 f.  
 rehabilitation, 21  
 Reider, Norman, xii  
 rejectees, 35 f., 40, 59-61, 118 f., 130-32, 336 (*see also* drop-outs)  
 relationships with patients, 207-208, 210  
     with people, 278-79  
 research, 61, 74, 128, 278, 308-309, 332-36, 338  
 residents, Menninger School of Psychiatry  
     after residency, 70-76  
     certification of, 71  
     changes in, during psychiatric education, 52-69  
     developmental styles, 52-61  
     dissatisfactions of, 74  
     entering, 33-51  
         early background, 37-40, 222-24, 281  
         expectations regarding future practice, 42-44  
         inter-class differences, 35-37  
         origins of interest in psychiatry, 44-51

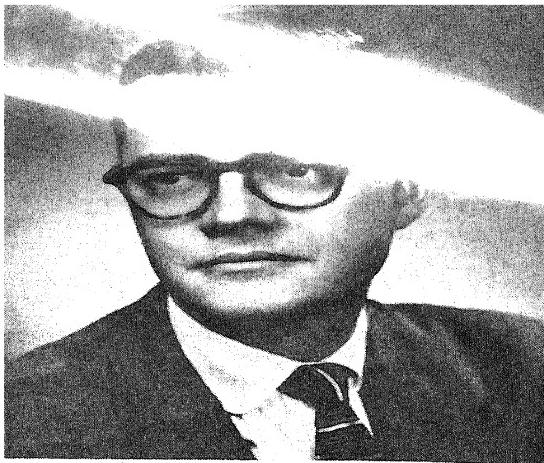
- residents (*cont.*)  
     religious background, 38 f.  
     research group, 34-35  
     job analysis for, 137, 139-40  
     kinds of practice, 73-75  
     personal psychiatric treatment, 63-69  
     personal qualities of, 140  
     psychoanalytic training for, 72  
     settings of practice, 72-73  
     what they learned, 61-63
- Riggs, Austen, 40
- Robbins, Lewis L., xi f.
- Rohles, Frederick H., Jr., 222, 335
- Rorschach test, xi, 5, 129, 141, 151, 156,  
     213 ff., 278 n., 302, 305, 307, 311 f.,  
     316 ff., 321 f., 324, 326, 330, 334,  
     348  
     errors in evaluating, 248  
     manual, 146 f., 200, 202 f.  
     scores as predictors, 230-31
- Rosen, Albert, 209
- Rottersman, William, xii
- "rough diamonds," 55
- Ruesch, Jurgen, 256, 259
- Sachs, Hanns, 301
- Sager, Clifford J., 214 n.
- Salmon, Thomas, 40
- Sarbin, T. R., 100, 103-104
- Saskatchewan, 354
- Schafer, Roy, xii, 99, 321, 324 n., 326  
     *quoted*, 247 n.
- schizoid disorganization of thought and  
     affect, 178, 235, 244, 246, 291, 326  
     defined, 145
- Schlichter, R. J., 84
- Scholastic Aptitude Test, 86
- security, feelings of, 161, 168, 275, 282,  
     319-20  
     defined, 144
- selection research  
     actuarial tradition, 79-89  
     clinical tradition, 89-99  
     clinical and/or statistical predictions,  
         99-107  
     theoretical issues in, 79-107  
         (*see also* Design I, Design II)
- self-administered inventories, 224-27
- self-concept questionnaire, 278
- self-confidence, 55-56, 163, 165, 247, 275,  
     277 f., 310  
     defined, 143
- Self-Interpretation (of TAT), 141, 152-  
     53, 170-74, 214, 308, 311 ff., 348  
     manual, 150 f., 200, 202  
     specimen analysis (Dr. Abbott), 170-74
- self-objectivity (insight), 46, 152, 165,  
     173, 262, 265, 274, 283, 288, 310-11  
     definition of, 143
- self-orientation, 143-44, 310-13
- sensitivity, 155, 166, 235 f., 248, 256, 265,  
     267, 277, 288, 289, 290, 322 f.  
     perceptual, defined, 145
- Seventh Medical Economics Survey, 75
- sexual adjustment, mature, 161, 168, 177,  
     270, 274, 281, 288, 292, 317-19  
     defined, 144
- Shakow, David, 265
- Shawnee County Guidance Clinic, 25
- shock treatment, 42, 74, 76, 105
- shortage of physicians, 11, 13, 354
- Siegal, Richard, xi
- significance, tests of, 360-61  
     of mean differences, 362-63
- Simoneit, M., 91, 95
- Small Sample Study, 47, 138, 140, 142,  
     153, 224, 226 f., 334 f.
- social adjustment with co-workers, 167,  
     273, 278 f., 288, 319-20  
     defined, 144
- social injustice, consciousness of, 241, 269,  
     274, 317  
     defined, 144
- Social Science Research Council, 85
- social work, psychiatric, 7 ff., 16, 20
- Southard School, 18, 25, 61, 125
- Soviet Union, 355
- Stalnaker, John M., 84 f.
- state hospitals, 72  
     reform in, 11-12
- statistical prediction (*see* predictions,  
     actuarial)
- statistical techniques, 86 ff.  
     explanation of, 359-64  
         (*see also* actuarial system)
- status-mindedness, 273, 289, 305 f.  
     defined, 143
- "steady highs," 53, 60, 76
- stereotypy of thought, freedom from, 166,  
     275 f., 289 f., 322  
     defined, 145
- Strong, Edward K., 40, 83, 225

- Strong Vocational Interest Blank (SVIB),  
5, 40-42, 83, 141, 146, 224-26,  
231 f., 348  
Lawyer key, 231 f.  
Psychiatrist A key, 225, 349  
Psychologist key, 225, 231 f.  
Stuit, Dewey B., 84, 222 f.  
subjects (*see* residents; rejectees)  
success as a criterion, 116, 120, 134  
supervisors' evaluations, 121, 122-24,  
125 ff., 133, 185, 200 ff., 205 f.,  
208 f., 212, 214, 217, 220, 225, 228,  
231, 235, 273, 279, 281  
Supplementary Face Sheet, 37 n., 150,  
214, 278  
SVIB (*see* Strong Vocational Interest  
Blank)  
Szondi test, 129, 213, 224
- TAT (*see* Thematic Apperception Test)  
Taylor, Frederick W., 80  
teaching, 10 f., 26-32, 74, 354 f.  
Terman, Lewis M., *quoted*, 228 n.  
tests, psychological, 40-41, 80-84; 93  
battery of, 87-88, 99, 302-303, 329, 348  
choice of, for Design II, 140-42, 151-54  
current use of, in selection, 4-5, 129,  
339-40  
diagnostic, 97-99  
performance (*see* situational)  
ratings of, 130-34  
recommendations on use of, in selec-  
tion, 302-26, 329-32, 334-35, 344  
scores, as predictors of psychiatric com-  
petence, 229-33  
situational, 82, 91, 97  
"trade," 81  
two-tailed and one-tailed, 361-62  
(*see also* ACE intelligence test; Army  
Alpha; Army Air Force tests; Army  
General Classification Test; Bern-  
reuter Personality Inventory; Bi-  
net; California Personality Inven-  
tory; E Scale; F Scale; Graduate  
Record Examination; Kuder Prefe-  
rence Test; Medical Aptitude  
Test; Medical College Admission  
Test; Miller Analogies Test  
(MAT); Minnesota Multiphasic  
Personality Inventory; Picture Re-
- action Test (PRT); question-  
naires; Rorschach; Scholastic Apti-  
tude Test; Self-Interpretation;  
Strong Vocational Interest Blank  
(SVIB); Szondi; Thematic Apper-  
ception Test (TAT); Verbal IQ;  
Wechsler-Bellevue Scale; Word  
Association)
- Tetzlaff, Glenna Webster, xi  
Texas (*see* University of Texas)  
Thematic Apperception Test (TAT), 5,  
93, 141, 146, 152, 155 ff., 214 f.,  
302, 304, 306, 308 f., 311 ff., 316 ff.,  
321 f., 324 f., 334, 347 f.  
manuals, 150-51, 156, 199, 200-202, 352  
prediction errors, 245-46  
specimen analysis (Dr. Abbott), 160-70,  
173, 179-81  
(*see also* Self-Interpretation)
- Thorndike, Robert L., 218  
time-and-motion studies, 80  
tolerance (*see* ethnocentric prejudice)  
Tomkins, Silvan S., 200  
Topeka Institute for Psychoanalysis, 64,  
72, 125, 339, 341 f.  
Topeka Psychoanalytic Society, 339  
Topeka State Hospital, 26, 34 f., 354  
Tucker, Anthony C., 225  
Turner, E. L., 14
- unemployment, problems of, 82  
"uneven workers," 58, 76  
"uninterested," the, 59 f.  
United States Employment Service, 82  
University of Cincinnati, 5 n.  
University of Colorado, 5 n.  
University of Kansas, 23, 26  
University of Michigan, 5 n., 41  
Project on the Selection of Clinical  
Psychologists (*see* Michigan Proj-  
ect)  
University of Minnesota, 5 n., 81-82  
University of Texas, 5 n.
- VA hospitals, reform in, 9-11  
VA Mental Hygiene Clinic, 25  
validities, of clinical predictions (rat-  
ings), 130-36, 194-96, 208-12, 216-  
33, 348 (*see also* interview; pre-  
dictors)

- Verbal IQ, associated with psychiatric competence, 228 ff., 277, 323 n., 331, 348
- verbalization, quality of (verbal facility), 173, 242, 268, 275, 277 f., 288, 323 defined, 145
- Veterans Administration, x, 9, 17, 19, 21, 34, 72 f., 109 (*see also* VA hospitals, VA Mental Hygiene Clinic)
- Viteles, M. S., 82
- vocational research, 83, 85
- voice, quality of, as source of error, 237 f.
- Wallin, Paul, 86
- War Office Selection Board, 95
- Ward, Clyde H., *quoted*, 351
- Ward, Mary Jane, 11
- warmth, emotional, 55, 155, 237 f., 242 f., 248, 262, 266, 271, 279, 289, 292 f., 306-307
- Warner, Korean, xi
- Washington University (St. Louis), 5 n.
- Wechsler, David, 40, 151, 227
- Wechsler Adult Intelligence Scale, 230, 331
- Wechsler-Bellevue Scale of Adult Intelli- gence, 40, 129, 142, 146, 151, 213, 227-30, 231, 277, 309, 324 f., 331, 348 (*see also* Verbal IQ)
- Weir, John R., 355
- Weiss, Fae Kaufman, xi
- Whitehorn, John, xi
- Winslow, Walker, 18 n., 35 n.
- Winter VA Hospital, 19, 23, 34, 125, 129, 193  
from January 1946 to July 1948, 19-22  
from 1948 to 1951, 22
- Medical Rehabilitation, 21
- Mental Hygiene Clinic, 64
- organizational structure, 25
- teaching methods, 26-32
- Wittmann, M. Phyllis, 105 f.
- Woodworth, R. S., 82
- Woodworth Personal Data Sheet, 82
- Word Association test, 129, 142, 146, 151, 213, 304, 307, 311
- work capacities, 237, 246, 266, 277-78, 309
- World War I, 8, 18, 80, 82, 337
- World War II, 3, 7-8, 13, 15, 85, 338, 353
- Worthington, Robert L., xi
- Yale Medical School, 66







**ROBERT R. HOLT**, Ph.D., is currently Professor of Psychology in the Graduate School of Arts and Science, New York University, and Director of N.Y.U.'s Research Center for Mental Health. At The Menninger Foundation, Dr. Holt held several positions in succession, including Senior Psychologist and Director of the Psychological Staff. He has published extensively in professional journals and is a contributor to six books.



**LESTER LUBORSKY**, Ph.D., is Senior Psychologist in the Research Department of The Menninger Foundation. Formerly with the Department of Psychology at the University of Illinois, Dr. Luborsky received his training as a psychoanalyst at the Topeka Institute for Psychoanalysis. He has had numerous papers published in psychological and psychiatric journals.

UNIVERSAL  
LIBRARY



116 928

UNIVERSAL  
LIBRARY